

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 04:29:55 : Search time 2346.74 Seconds
(without alignments)
17786.465 Million cell updates/sec

Title: US-10-030-390-3

Perfect score: 8241

Sequence: 1 agatctaacatccaagacg.....aaaataacagttattatcg 8241

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3337386 seqs, 2532474682 residues

Total number of hits satisfying chosen parameters: 6674772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	6561.6	79.6	9276	10	US-09-776-191-45
2	6561.6	79.6	9276	15	US-10-190-0308-13
3	6561.6	79.6	9276	15	US-10-302-840A-13
4	6561.6	79.6	9276	15	US-10-287-219-13
5	6561.6	79.6	9276	15	US-10-112-221A-11
6	6561.6	79.6	9276	15	US-10-104-271-11
7	6561.6	79.6	9276	16	US-10-147-211A-13
8	6383.8	77.5	9776	10	US-09-883-848A-31
9	5709.8	69.3	10462	10	US-09-883-848A-34
10	5665.8	68.8	10491	10	US-09-883-848A-32
11	5657.8	68.7	10512	10	US-09-883-848A-33
12	5142.8	62.4	8584	15	US-10-038-722-71
13	5130.8	62.3	8590	15	US-10-038-722-75
14	4565.2	55.4	8157	15	US-10-038-722-70

C 15 2659 32.3 2659 15 US-10-179-046-19 Sequence 19, Appl
C 16 2404.6 29.2 5740 10 US-09-921-380-3 Sequence 3, Appl
C 17 2320.8 28.2 5934 17 US-10-612-410-2 Sequence 2, Appl
C 18 2320.8 28.2 7618 17 US-10-612-410-1 Sequence 1, Appl
C 19 2314 28.1 11918 17 US-10-612-224-2 Sequence 2, Appl
C 20 2314 28.1 13227 17 US-10-612-224-4 Sequence 4, Appl
C 21 2314 28.1 13278 17 US-10-612-224-3 Sequence 3, Appl
C 22 2312.4 28.1 11466 17 US-10-466-959-4 Sequence 4, Appl
C 23 2302 27.9 4899 17 US-10-600-790-19 Sequence 19, Appl
C 24 2300.4 27.9 6865 15 US-10-338-411-33 Sequence 33, Appl
C 25 2300.4 27.9 6865 16 US-10-389-640-33 Sequence 33, Appl
C 26 2300.4 27.9 7894 15 US-10-338-411-34 Sequence 34, Appl
C 27 2300.4 27.9 7894 16 US-10-389-640-34 Sequence 34, Appl
C 28 2298.6 27.9 11152 9 US-09-847-101B-12 Sequence 12, Appl
C 29 2298.6 27.9 11152 10 US-09-482-682-12 Sequence 12, Appl
C 30 2298.8 27.9 8068 15 US-10-153-006-27 Sequence 27, Appl
C 31 2298.2 27.9 7617 13 US-10-159-006-35 Sequence 35, Appl
C 32 2298.2 27.9 7626 13 US-10-401-000-1 Sequence 1, Appl
C 33 2298.2 27.9 9511 9 US-10-401-000-2 Sequence 2, Appl
C 34 2298.2 27.9 9511 9 US-09-897-006-34 Sequence 34, Appl
C 35 2298.2 27.9 9511 10 US-09-897-511A-34 Sequence 34, Appl
C 36 2298.2 27.9 9511 13 US-10-397-079-34 Sequence 34, Appl
C 37 2297.8 27.9 8685 14 US-10-098-035-3 Sequence 3, Appl
C 38 2297.4 27.9 8687 17 US-10-467-546-15 Sequence 15, Appl
C 39 2297.4 27.9 9362 17 US-10-467-546-16 Sequence 16, Appl
C 40 2297.4 27.9 9400 17 US-10-467-546-17 Sequence 17, Appl
C 41 2297.2 27.9 11641 13 US-10-684-300-13 Sequence 13, Appl
C 42 2297.2 27.9 11641 13 US-09-993-059-33 Sequence 33, Appl
C 43 2297.2 27.9 11641 15 US-10-103-327-33 Sequence 33, Appl
C 44 2297.2 27.9 11641 17 US-10-684-349-13 Sequence 13, Appl

ALIGNMENTS

RESULT 1

US-09-776-191-45
; Sequence 45, Application US/09776191
; Publication No. US20030119168A1
; GENERAL INFORMATION:
; APPLICANT: Edwin L. Madison
; APPLICANT: Edgar O. Ong
; APPLICANT: Giann-Chern Yeh
; APPLICANT: Corvas International, Inc.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: TRANSMEMBRANE SERINE PROTEASES, THE ENCODED PROTEINS AND
; TITLE OF INVENTION: METHODS BASED THEREON
; FILE REFERENCE: 24745-1607
; CURRENT APPLICATION NUMBER: US/09/776,191
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/213,124
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/234,840
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/179,982
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 60/183,542
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: 09/657,968
; PRIOR FILING DATE: 2000-02-08
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 9276
; TYPE: DNA
; ORGANISM: Pichia pastoris
US-09-776-191-45

Query Match 79.6%; Score 6561.6; DB 10; Length 9276;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;

QY 1 AGATCTAACATCCAAAGAGGAGTTGATGAACCTTTTGGCATCCGACATCCACAG 60
 Db |||||
 QY 1 AGATCTAACATCCAAAGAGGAGTTGATGAACCTTTTGGCATCCGACATCCACAG 60
 Db |||||
 QY 61 GTCCATTCTCACATAGTCCAAACGCAACAGAGGGGATACACTAGCAGCAGACCGT 120
 Db |||||
 QY 61 GTCCATTCTCACATAGTCCAAACGCAACAGAGGGGATACACTAGCAGCAGACCGT 120
 Db |||||
 QY 121 TGCAAACGAGGAGCTCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
 Db |||||
 QY 121 TGCAAACGAGGAGCTCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
 Db |||||
 QY 181 AGCCAGGTTATGGGCTTGATGGAGCTCGCTCATTCCTCTCTCTCTCTCTCTCTCT 240
 Db |||||
 QY 181 AGCCAGGTTATGGGCTTGATGGAGCTCGCTCATTCCTCTCTCTCTCTCTCTCTCT 240
 Db |||||
 QY 241 ACACCATGACTTTATAGCTGTCTATCTGGGCCCCCTGGCGAGGTTTCATGTTGTTTA 300
 Db |||||
 QY 241 ACACCATGACTTTATAGCTGTCTATCTGGGCCCCCTGGCGAGGTTTCATGTTGTTTA 300
 Db |||||
 QY 301 TTTCGGAATGCAACAGCTCCGATACACCCGAAACATCACTCCAGATAGGGCTTCTG 360
 Db |||||
 QY 301 TTTCGGAATGCAACAGCTCCGATACACCCGAAACATCACTCCAGATAGGGCTTCTG 360
 Db |||||
 QY 361 AGTGTGGGCTCAAAATAGTTTCATGTTCCCAAAATGGCCCAAACTGACAGCTTAAACGCT 420
 Db |||||
 QY 361 AGTGTGGGCTCAAAATAGTTTCATGTTCCCAAAATGGCCCAAACTGACAGCTTAAACGCT 420
 Db |||||
 QY 421 GTCTTGGAACTTAATGACAAAAGCGTGATCTCATCCAGATGAACCTAAGTTGGTTTCG 480
 Db |||||
 QY 421 GTCTTGGAACTTAATGACAAAAGCGTGATCTCATCCAGATGAACCTAAGTTGGTTTCG 480
 Db |||||
 QY 481 TTGAAATGCTAACGCGCAGTTGGTCAAAAGAACTTCCAAAGTCCGCAATACCGTTGT 540
 Db |||||
 QY 481 TTGAAATGCTAACGCGCAGTTGGTCAAAAGAACTTCCAAAGTCCGCAATACCGTTGT 540
 Db |||||
 QY 541 CTGTTTGGTATGATGACGAATGCTCAAAATAAATCTCAATTAATGTTAGCGCAGTCT 600
 Db |||||
 QY 541 CTGTTTGGTATGATGACGAATGCTCAAAATAAATCTCAATTAATGTTAGCGCAGTCT 600
 Db |||||
 QY 601 CTCATATCGTCTGAAACCCCGTGACCTGTGCGGAAACGCAATGGGAAACACCCGCT 660
 Db |||||
 QY 601 CTCATATCGTCTGAAACCCCGTGACCTGTGCGGAAACGCAATGGGAAACACCCGCT 660
 Db |||||
 QY 661 TTTTGGATGATATGCAATGTCTCCACATTTGATGCTTCCAAAGTCTGTTGGGAAATACT 720
 Db |||||
 QY 661 TTTTGGATGATATGCAATGTCTCCACATTTGATGCTTCCAAAGTCTGTTGGGAAATACT 720
 Db |||||
 QY 721 GCTGATAGCTTAACGTTTATGATCAAAATTTAACTGTTCTAACCCCTACTTGACAGCAAT 780
 Db |||||
 QY 721 GCTGATAGCTTAACGTTTATGATCAAAATTTAACTGTTCTAACCCCTACTTGACAGCAAT 780
 Db |||||
 QY 781 ATATAAAGAGGAGGAGTCCCTGTCTTAAACCTTTTATCATCATTTATAGCTT 840
 Db |||||
 QY 781 ATATAAAGAGGAGGAGTCCCTGTCTTAAACCTTTTATCATCATTTATAGCTT 840
 Db |||||
 QY 841 ACTTTTCATATTCGACCTGGTTTCCAAATTTGACAGCTTTGATTTTAAAGCTTTTAAACGA 900
 Db |||||
 QY 841 ACTTTTCATATTCGACCTGGTTTCCAAATTTGACAGCTTTGATTTTAAAGCTTTTAAACGA 900
 Db |||||
 QY 901 CAACCTTGAGAGATCAAAACCAACTAATTTTCGAGAGATCCAAACGATGAGATTTCCCT 960
 Db |||||
 QY 901 CAACCTTGAGAGATCAAAACCAACTAATTTTCGAGAGATCCAAACGATGAGATTTCCCT 960
 Db |||||
 QY 961 TCAATTTTACTGCGAGTTTATTCGAGATCTCCCGCATCTCCCGCATTTAGCTGCTCCAGTCAACACT 1020
 Db |||||
 QY 961 TCAATTTTACTGCGAGTTTATTCGAGATCTCCCGCATCTCCCGCATTTAGCTGCTCCAGTCAACACT 1020
 Db |||||
 QY 1021 ACAACAGAGATGAACGCAAAATTCGGCTGAAGCTGTATCTCGGTTTACTTCAGATTTTA 1080
 Db |||||
 QY 1021 ACAACAGAGATGAACGCAAAATTCGGCTGAAGCTGTATCTCGGTTTACTTCAGATTTTA 1080
 Db |||||
 QY 1081 GAAGGGGATTTCCGATGTTGCTGTTTGGCCATTTTCCAAACAGCAAAATAACGGGTTATTG 1140
 Db |||||

Db 1081 GAAGGGGATTTCCGATGTTGCTGTTTGGCCATTTTCCAAACAGCAAAATAACGGGTTATTG 1140
 QY |||||
 Db 1141 TTTTATAAATACTACTATTTCAGCATTTCTCTAAAGAAAGGGGTATCTCTCGAGAAA 1200
 QY |||||
 Db 1141 TTTTATAAATACTACTATTTCAGCATTTCTCTAAAGAAAGGGGTATCTCTCGAGAAA 1200
 QY |||||
 Db 1201 AGAGAGGCTGAAGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCC 1260
 QY |||||
 Db 1201 AGAGAGGCTGAAGCTTTA----- 1217
 QY |||||
 Db 1261 GCCCCCAGGAGAGGATAAATTTGGGCTTCCCCGGTGTCCACGCCAGCAGTGCACGGAG 1320
 QY |||||
 Db 1218 ----- 1217
 QY |||||
 Db 1321 AGAGGTTGCTGTTTGTGATGACAGTGTCCGGGATTCGGTGGTCTCCACCCCATGGCC 1380
 QY |||||
 Db 1218 ----- 1217
 QY |||||
 Db 1381 ATCGAGAACTCAAGAAAGAAATGTCCCTTCTAACTAGTGGGTAGAAATCCCTTAGGG 1440
 QY |||||
 Db 1218 -----CGTAGAATTCCTTAGGG 1234
 QY |||||
 Db 1441 CGGCCGGAATTAATTCGCTTGAACATGACTGTTCTCAGTTTCAAGTTGGGCACTTACG 1500
 QY |||||
 Db 1235 CGGCCGGAATTAATTCGCTTGAACATGACTGTTCTCAGTTTCAAGTTGGGCACTTACG 1294
 QY |||||
 Db 1501 AGAAGACCGGCTCTGCTAGATTCTAATCAAGAGGATGTCAAGATGCCATTTGCCCTGAGAG 1560
 QY |||||
 Db 1295 AGAAGACCGGCTCTGCTAGATTCTAATCAAGAGGATGTCAAGATGCCATTTGCCCTGAGAG 1354
 QY |||||
 Db 1561 ATGCGAGCTTCAATTTTGAATCTTTTATTTGTAACCTATATAGTATAGGATTTTCTT 1620
 QY |||||
 Db 1355 ATGCGAGCTTCAATTTTGAATCTTTTATTTGTAACCTATATAGTATAGGATTTTCTT 1414
 QY |||||
 Db 1621 GTCATTTCTTCTCTGACAGCTTCTCTCCTCATCAGCCTATCTCGCAGCTGATGAA 1680
 QY |||||
 Db 1415 GTCATTTCTTCTCTGACAGCTTCTCTCCTCATCAGCCTATCTCGCAGCTGATGAA 1474
 QY |||||
 Db 1681 TATCTTGGTATAGGGTTTGGGAAATCATTCGAGTTTGTATGTTTCTTGGTATTCTCC 1740
 QY |||||
 Db 1475 TATCTTGGTATAGGGTTTGGGAAATCATTCGAGTTTGTATGTTTCTTGGTATTCTCC 1534
 QY |||||
 Db 1741 ACTCTCTCTTCAGAGTACAGAGATTAAGTGAAGTTCGTTTGTCAAGCTTATCGATAA 1800
 QY |||||
 Db 1535 ACTCTCTCTTCAGAGTACAGAGATTAAGTGAAGTTCGTTTGTCAAGCTTATCGATAA 1594
 QY |||||
 Db 1801 GCTTTAATCGGTAGTTTATCAAGTTAAATTCGTTAAGCAGTACAGGACCGGTGATGAA 1860
 QY |||||
 Db 1595 GCTTTAATCGGTAGTTTATCAAGTTAAATTCGTTAAGCAGTACAGGACCGGTGATGAA 1654
 QY |||||
 Db 1861 ATCTAACATGCGCTCATGCTCCTCGGACCGCTCCTCGGATGCTGTAGGATAGG 1920
 QY |||||
 Db 1655 ATCTAACATGCGCTCATGCTCCTCGGACCGCTCCTCGGATGCTGTAGGATAGG 1714
 QY |||||
 Db 1921 CTGTTGTTATGCGGTACTCCCGGGCTCTTTCGGGATATCGTCCATTCGACAGATCGC 1980
 QY |||||
 Db 1715 CTGTTGTTATGCGGTACTCCCGGGCTCTTTCGGGATATCGTCCATTCGACAGATCGC 1774
 QY |||||
 Db 1981 CAGTCACTATGGGCTGCTAGCGGTATATGCGGTGATGCAATTTCTATGCGCACCGCT 2040
 QY |||||
 Db 1775 CAGTCACTATGGGCTGCTAGCGGTATATGCGGTGATGCAATTTCTATGCGCACCGCT 1834
 QY |||||
 Db 2041 TCTCGGAGACACTGTCCGACCGCTTTGGCCGCGCCAGCTCTGCTCGCTTCTGCTACTGG 2100
 QY |||||
 Db 1835 TCTCGGAGACACTGTCCGACCGCTTTGGCCGCGCCAGCTCTGCTCGCTTCTGCTACTGG 1894
 QY |||||
 Db 2101 AGCCACTATCGACTAGCGCATCATGGGACACACCCGCTCTGTGATCTATCGATCTA 2160
 QY |||||
 Db 1895 AGCCACTATCGACTAGCGCATCATGGGACACACCCGCTCTGTGATCTATCGATCTA 1954
 QY |||||
 Db 2161 AATGTAAGTTAAATCTCTAAATTAATTAATTAAGTCCCAAGTTTCTCCATFACGACCTTAA 2220
 QY |||||

Db 1955 AATGTAAGTTAAATCTCTAAATAATAAATAAGTCCAGTTTCTCCATACGAACCTTAA 2014
Qy
Qy 2221 CAGCATTTGGGTGAGCATCTAGACCTTCAAAGCAGCCAGATCCATCACTCTTGGCCAA 2280
Db 2015 CAGCAATTCGGGTGAGCATCTAGACCTTCAAAGCAGCCAGATCCATCACTCTTGGCCAA 2074
Qy 2281 TATGTTTCAGTCCCTCAGGAGTTAGCTCTTGTGAAGTGAATGAATCTTCTGGAAGGTTGCAG 2340
Db 2075 TATGTTTCAGTCCCTCAGGAGTTAGCTCTTGTGAAGTGAATGAATCTTCTGGAAGGTTGCAG 2134
Qy 2341 TGTTAACTCCCGTGTATTTAGCGGGCATATCCGTAGCTTGGAAAAGTGTGTTGTACCGG 2400
Db 2135 TGTTAACTCCCGTGTATTTAGCGGGCATATCCGTAGCTTGGAAAAGTGTGTTGTACCGG 2194
Qy 2401 AGGAGTAATCTCCAACTCTCTGAGAGTAGGCAACAAACACACAGATCCACGCTGTT 2460
Db 2195 AGGAGTAATCTCCAACTCTCTGAGAGTAGGCAACAAACACACAGATCCACGCTGTT 2254
Qy 2461 GTACTTGATCAACATPAGAAGAGATTTCTCGATTTGAGGATCAAGTGTTCAGGAGCGT 2520
Db 2255 GTACTTGATCAACATPAGAAGAGATTTCTCGATTTGAGGATCAAGTGTTCAGGAGCGT 2314
Qy 2521 ACTGATTTGGACATTTCCAAAGCCCTGCTGTAAGTTGCAACGATAGGTTGTAGAGTGTG 2580
Db 2315 ACTGATTTGGACATTTCCAAAGCCCTGCTGTAAGTTGCAACGATAGGTTGTAGAGTGTG 2374
Qy 2581 CAATACACTTCGCTACAAATTTCAACCCCTTGCAACTGACAGCTTGGTTGTGAACAGCAT 2640
Db 2375 CAATACACTTCGCTACAAATTTCAACCCCTTGCAACTGACAGCTTGGTTGTGAACAGCAT 2434
Qy 2641 CTTCAATTTCTGGCAGCTCTTGTCTGTCAATATCGACAGCCACAGATCACTCTGGGAAT 2700
Db 2435 CTTCAATTTCTGGCAGCTCTTGTCTGTCAATATCGACAGCCACAGATCACTCTGGGAAT 2494
Qy 2701 CAATACCAATTTGAGTACAGAGGCTCTGAGGCAACGAAATCTGGATCAAGCTGATTT 2760
Db 2495 CAATACCAATTTGAGTACAGAGGCTCTGAGGCAACGAAATCTGGATCAAGCTGATTT 2554
Qy 2761 TATCAGCAATTAATAGAACTTTAGAAAGCCAGAGGATGTCAATATCTACAGAGGCTG 2820
Db 2555 TATCAGCAATTAATAGAACTTTAGAAAGCCAGAGGATGTCAATATCTACAGAGGCTG 2614
Qy 2821 ATGTGTCATTTGTAACCATCATCTTGGCAGCAGTAAAGAACTGTTTCTGGACCAATA 2880
Db 2615 ATGTGTCATTTGTAACCATCATCTTGGCAGCAGTAAAGAACTGTTTCTGGACCAATA 2674
Qy 2881 TTTTGTCACTTAGGAACAGTTTCTGTTCCGTAAAGCCATAGCAGCTACTGCTGGGCGC 2940
Db 2675 TTTTGTCACTTAGGAACAGTTTCTGTTCCGTAAAGCCATAGCAGCTACTGCTGGGCGC 2734
Qy 2941 CTCCTGCTAGCAGATACACTTAGACCAACCTTGTGGGCAACGTAGATGACTTCTGGGG 3000
Db 2735 CTCCTGCTAGCAGATACACTTAGACCAACCTTGTGGGCAACGTAGATGACTTCTGGGG 2794
Qy 3001 TAAAGGTACCATCTCTTTAGGTGAGATGCAAAAACAAATTTCTTTGCAACCAACCAACTT 3060
Db 2795 TAAAGGTACCATCTCTTTAGGTGAGATGCAAAAACAAATTTCTTTGCAACCAACCAACTT 2854
Qy 3061 TGGCAGGAACACCCAGCATCAGGAAGTGGAAAGGAGAAATTTGCGGTTCCACAGGAATAT 3120
Db 2855 TGGCAGGAACACCCAGCATCAGGAAGTGGAAAGGAGAAATTTGCGGTTCCACAGGAATAT 2914
Qy 3121 AGAGGCCAACTTTCTCAATAGTCTTGCAAAACAGAGCAGACTACACAGGGCAAGTCT 3180
Db 2915 AGAGGCCAACTTTCTCAATAGTCTTGCAAAACAGAGCAGACTACACAGGGCAAGTCT 2974
Qy 3181 CAACTTGCAACGCTCTCGTTAGTTAGCTTCAATGAAATTTCTGACGTTATCTATAGAGA 3240
Db 2975 CAACTTGCAACGCTCTCGTTAGTTAGTTCAATGAAATTTCTGACGTTATCTATAGAGA 3034
Qy 3241 GATCAATGGCTCTTACAGTTATCTGCAATTCATAGTTTCTCTGGGAAGGAGCTT 3300
Db 3035 GATCAATGGCTCTTACAGTTATCTGCAATTCATAGTTTCTCTGGGAAGGAGGCTT 3094

Qy 3301 CTAACACAGGTCTCTTCAAGCGACTCCATCAAACTTGGCAGTTAGTTCTTAAAGGCGTT 3360
Db 3095 CTAACACAGGTCTCTTCAAGCGACTCCATCAAACTTGGCAGTTAGTTCTTAAAGGCGTT 3154
Qy 3361 TGTCAACATTTTGACGAACAATTTGTGACAAATTTGGTTTGAATAATTCATTAATCTGTTCG 3420
Db 3155 TGTCAACATTTTGACGAACAATTTGTGACAAATTTGGTTTGAATAATTCATTAATCTGTTCG 3214
Qy 3421 TTTTCTGGATAGGACGACGAAGGGCATCTTCAATTTCTTGTGAGGAGGCTTAGAAACGT 3480
Db 3215 TTTTCTGGATAGGACGACGAAGGGCATCTTCAATTTCTTGTGAGGAGGCTTAGAAACGT 3274
Qy 3481 CAATTTTGACACAATTTCAATACGACCTTCAAGAGGAGCTTCTTTAGGTTTGGATTTCTTCT 3540
Db 3275 CAATTTTGACACAATTTCAATACGACCTTCAAGAGGAGCTTCTTTAGGTTTGGATTTCTTCT 3334
Qy 3541 TAGGTTTGTCTTGGTGTATCTTGGCTTGGCATCTCTCTTCTCTCTCTCTCTCTCTCTCTCT 3600
Db 3335 TAGGTTTGTCTTGGTGTATCTTGGCTTGGCATCTCTCTTCTCTCTCTCTCTCTCTCTCTCT 3394
Qy 3601 ACTTCATATCCAGGTTTCTCTGACCTCGTCAAGCTCAACGCTACTTGGCAGCATCTAA 3660
Db 3395 ACTTCATATCCAGGTTTCTCTGACCTCGTCAACGCTCAACGCTACTTGGCAGCATCTAA 3454
Qy 3661 CTAATCCAAAATAAAATAAGTCAGCAATTTCCAGGCTATATCTTCTCTCTCTCTCTCTCTCT 3720
Db 3455 CTAATCCAAAATAAAATAAGTCAGCAATTTCCAGGCTATATCTTCTCTCTCTCTCTCTCTCT 3514
Qy 3721 CTGCAAGTTTCATCAGCTTCT 3780
Db 3515 CTGCAAGTTTCATCAGCTTCT 3574
Qy 3781 ACCGTTTGGTATTAAGAACTTCTTGGAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3840
Db 3575 ACCGTTTGGTATTAAGAACTTCTTGGAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3634
Qy 3841 TGGCTCTAAGACCTTTGATTTGGCCAAAACAGGAGTGGCTTCCAAAGTGCAGAGAAACCA 3900
Db 3635 TGGCTCTAAGACCTTTGATTTGGCCAAAACAGGAGTGGCTTCCAAAGTGCAGAGAAACCA 3694
Qy 3901 CACTCTTTTGTTCACCAAAAATTTCAAGCAGTCTCCATCACAATTCGAATCCCA 3960
Db 3695 CACTCTTTTGTTCACCAAAAATTTCAAGCAGTCTCCATCACAATTCGAATCCCA 3754
Qy 3961 GCAACTTTTGTAGTGTCTCAGATGTAGACCTTTTATACCAAAACGCTGACAGAGAT 4020
Db 3755 GCAACTTTTGTAGTGTCTCAGATGTAGACCTTTTATACCAAAACGCTGACAGAGAT 3814
Qy 4021 GGTAGACTCCAGTTTGTCTTATAGCTCCGGAATAGACTTTTGGACGAGTACACCA 4080
Db 3815 GGTAGACTCCAGTTTGTCTTATAGCTCCGGAATAGACTTTTGGACGAGTACACCA 3874
Qy 4081 GGCCCAACGAGTAATTTAGAAGAGTCCAGCCACCAAAAGTGTGAATAGACCATCGGGCGGT 4140
Db 3875 GGCCCAACGAGTAATTTAGAAGAGTCCAGCCACCAAAAGTGTGAATAGACCATCGGGCGGT 3934
Qy 4141 CAGTAGTCAAGACGCCCAAAAATTTCACTGACAGGGAATTTTGTGACATCTTCTAGAAA 4200
Db 3935 CAGTAGTCAAGACGCCCAAAAATTTCACTGACAGGGAATTTTGTGACATCTTCTAGAAA 3994
Qy 4201 GTTTCGTATTTCACTAGTCAATTTCCAGCATCAATTAATGGGATTTATACCAAGCAACAG 4260
Db 3995 GTTTCGTATTTCACTAGTCAATTTCCAGCATCAATTAATGGGATTTATACCAAGCAACAG 4054
Qy 4261 TGGAGTCACTCTACCACTTTTGGGCTCTAGAAAAAGCATTAACAGTTTCTACTCCGC 4320
Db 4055 TGGAGTCACTCTACCACTTTTGGGCTCTAGAAAAAGCATTAACAGTTTCTACTCCGC 4114
Qy 4321 CATTTGTGAACTTTTCAATCGCAGTGGAGAGAAAAAGGCAAGCATCTAGCAT 4380
Db 4115 CATTTGTGAACTTTTCAATCGCAGTGGAGAGAAAAAGGCAAGCATCTAGCAT 4174

QY 4381 TAGCGGCAAGGATGCAACTTTTATCAACAGGCTCTTATAGATAAACCTTAGCGCTGGGA 4440
DB 4175 TAGCGGCAAGGATGCAACTTTTATCAACAGGCTCTTATAGATAAACCTTAGCGCTGGGA 4234
QY 4441 TCATCCTTTGGACACACTTTTCTGCCAANTCTAGTCCCAAAATCACTTCAATGATACCAT 4500
DB 4235 TCATCCTTTGGACACACTTTTCTGCCAANTCTAGTCCCAAAATCACTTCAATGATACCAT 4294
QY 4501 TATTGTACAACTTGAGCAAGTGTGCGATCAGCTCCTCAAAATTTGCTCTCTGTAAACGATG 4560
DB 4295 TATTGTACAACTTGAGCAAGTGTGCGATCAGCTCCTCAAAATTTGCTCTCTGTAAACGATG 4354
QY 4561 ACTCAACTTGCACATTAACCTTGAAGCTCAGTCAATGAGTGAACCTTGAATGAGTGTGCA 4620
DB 4355 ACTCAACTTGCACATTAACCTTGAAGCTCAGTCAATGAGTGAACCTTGAATGAGTGTGCA 4414
QY 4621 GCTGCTCAGCAGCATAGGAAAACAGCGCTTTTCTACCAAACTCAAGGAATTTACAACT 4680
DB 4415 GCTGCTCAGCAGCATAGGAAAACAGCGCTTTTCTACCAAACTCAAGGAATTTACAACT 4474
QY 4681 CTGCAACAATTGCGTATGCAAGTAGCAAGGAAATGTCATCTTGAAGTCGACAGTGAG 4740
DB 4475 CTGCAACAATTGCGTATGCAAGTAGCAAGGAAATGTCATCTTGAAGTCGACAGTGAG 4534
QY 4741 TGTAGTCTTGAGAAATTTCTGAGCGGTATTTTATATCATGAGTCAGTCATCAGGAGA 4800
DB 4535 TGTAGTCTTGAGAAATTTCTGAGCGGTATTTTATATCATGAGTCAGTCATCAGGAGA 4594
QY 4801 TCCTCTACCGCGGACGATCGTGCCGA----- 4828
DB 4595 TCCTCTACCGCGGACGATCGTGCCGAACCTGACAGGGGGGGGGCGCTGAGCTGCG 4654
QY 4829 ----- 4828
DB 4655 CTCGTGAAGAGGTGTGCTGACTCATACAGGCTGAAATCGGCCCATATCAGCAGCAGA 4714
QY 4829 ----- 4828
DB 4715 AAGTGAGGAGCCAGGTGTGATGAGAGCTTTGTGTAGTGGACCAAGTTGGTGAATTTGA 4774
QY 4829 ----- 4828
DB 4775 ACTTTTGCTTTGCCACGGAACGGTCTGCGTTGTCGGAAGATGCGTGATCTGATCCTTCA 4834
QY 4829 ----- 4828
DB 4835 ACTCAGCAAAAGTTGATTTATTCAACAAGCGCGCTCCGTCAGTCAGCGTAATGCT 4894
QY 4829 ----- 4828
DB 4895 CTGCCAGTGTTAACAACCAATTAACCAATTTCTGATTAGAAAACTCATCGAGCATCAATG 4954
QY 4829 ----- 4828
DB 4955 AAACGTCAANTTTATCATATCAGGATTAATCAATACCATATTTTGAAGAAAGCGTTTCG 5014
QY 4829 ----- 4828
DB 5015 TAATGAAGGAGAAAACCTACCGAGGACAGTTCCATAGGATGCAAGATCCTGTTATCGCTC 5074
QY 4829 ----- 4828
DB 5075 TCGGATTCGAGCTCGTCCAAATCAATACCACTTAATTTTCCCTCGCTCAAAAATAAG 5134
QY 4829 ----- 4828
DB 5135 GTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAAATGCAAAAGCTT 5194
QY 4829 ----- 4828
DB 5195 ATGCATTTCTTCCAGACTTGTTCACAGGCGCAGCCATTAACGCTCGTCATCAAAATCACT 5254
QY 4829 ----- 4828

DB 5255 CGCATCAACCAACCGTTATTTCATTGTTGGCTTGGAGAGAGCAATACGCGATC 5314
QY 4829 ----- 4828
DB 5315 GCTGTTAAAGGACAATTAACAACAGGAATCGAATGCAACCGCGCAGGAACACTGCCAG 5374
QY 4829 ----- 4828
DB 5375 CGCATCAACCAATATTTCACCTGAATCAGGATATTCTTCTAATACCTGGAAATGCTGTTTT 5434
QY 4829 ----- 4828
DB 5435 CCCGGGATGCAAGTGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTTGAT 5494
QY 4829 ----- 4828
DB 5495 GGTCCGAAGAGGCATAAATTCCTGTCAGCCAGTTTGTCTGTGACCATCTCATCTGTAAACATC 5554
QY 4829 ----- 4828
DB 5555 ATTGCAACGCTACCTTTTGCCATGTTTCAGAAACAACACTCTGGCGCATCGGCTTCCCAT 5614
QY 4829 ----- 4828
DB 5615 CAATCGATAGATTGTCCGACCTGATTGCGGACATTATCGGAGCCCATTTATACCCATA 5674
QY 4829 ----- 4828
DB 5675 TAAATCAGCATCCATGTTGGAATTTAATCGCGGCTCGAGCAAGACGTTTCCCGTTGAAT 5734
QY 4829 ----- 4828
DB 5735 ATGGCTCATACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTATTGTTCAAGA 5794
QY 4829 ----- 4828
DB 5795 TGATATATTTTATCTTGTGCAATGTAAATCATCAGAGATTTTGAGACACAACGTGGCTTC 5854
QY 4829 -----CCTGCAAGTCCGCAATCACCGGCGCCACAGGTCGGTGTGCTGGCGCTATAT 4879
DB 5855 CCCCCCCCCCTGCAAGTCCGCAATCACCGGCGCCACAGGTCGGTGTGCTGGCGCTATAT 5914
QY 4880 CGCCGACATCACGATGCGGAAGATCGGGCTCGCACTTCGGGCTCATGAGCGCTTGT 4939
DB 5915 CGCCGACATCACGATGCGGAAGATCGGGCTCGCACTTCGGGCTCATGAGCGCTTGT 5974
QY 4940 CGGCGTGGGTATGTTGGCAGGCCCCCGTGGCGGGGACTGTTGGCGGCAATCTCTTGTGA 4999
DB 5975 CGGCGTGGGTATGTTGGCAGGCCCCCGTGGCGGGGACTGTTGGCGGCAATCTCTTGTGA 6034
QY 5000 TGCAACAATCTTGGCGGCGGCTGCTCAACGGCTCAACCTACTACTTGGGCTGCTTCT 5059
DB 6035 TGCAACAATCTTGGCGGCGGCTGCTCAACGGCTCAACCTACTACTTGGGCTGCTTCT 6094
QY 5060 AATGAGGAGTCCGATTAAGGAGAGCGTCGAGTATCTATGATTGGAAGTATGGAAATGCT 5119
DB 6095 AATGAGGAGTCCGATTAAGGAGAGCGTCGAGTATCTATGATTGGAAGTATGGAAATGCT 6154
QY 5120 GATACCCGATCTTTCAGTGTCTTGAAGTCTCTATCAGATATGCCCCAATTAAGCAAC 5179
DB 6155 GATACCCGATCTTTCAGTGTCTTGAAGTCTCTATCAGATATGCCCCAATTAAGCAAC 6214
QY 5180 CGGAGGAGGAGATTTTCAGTAAATTTCTCTGACTTTTGGTCACTAGTAGACTCGAATG 5239
DB 6215 CGGAGGAGGAGATTTTCAGTAAATTTCTCTGACTTTTGGTCACTAGTAGACTCGAATG 6274
QY 5240 TGAGACTATCTCGGTTATGACAGCAGAAATGTCCTTTTGGAGACAGTAAATGAGTCCC 5299
DB 6275 TGAGACTATCTCGGTTATGACAGCAGAAATGTCCTTTTGGAGACAGTAAATGAGTCCC 6334
QY 5300 ACCAATAAGAAATCTCTGTTATCAGGACAACTCTTGTTCGAACTTTTTCGCTGCC 5359

QY 7520 TCAGAGTAAAGTGGCCGAGTGTATCACTCATGTTATGGCAGCACTGCATATTTCTC 7579
Db |||||
QY 8555 TCAGAGTAAAGTGGCCGAGTGTATCACTCATGTTATGGCAGCACTGCATATTTCTC 8614
Db |||||
QY 7580 TTAAGTCTATGATCCGATCCGTAAGATGCTTTTCTGTGACCTGGTGGTCAACCAAGTCAT 7639
Db |||||
QY 8615 TTAAGTCTATGATCCGTAAGATGCTTTTCTGTGACCTGGTGGTCAACCAAGTCAT 8674
Db |||||
QY 7640 TCTGAGATAGTCTATGCGGACCGAGTGTGCTCTTGGCCGCGCTCAACACGCGATATA 7699
Db |||||
QY 8675 TCTGAGATAGTCTATGCGGACCGAGTGTGCTCTTGGCCGCGCTCAACACGCGATATA 8734
Db |||||
QY 7700 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCTTGGGAAACGTTCTTCGGGGCGAA 7759
Db |||||
QY 8735 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCTTGGGAAACGTTCTTCGGGGCGAA 8794
Db |||||
QY 7760 AACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATCAACCCACTCGTCACCCA 7819
Db |||||
QY 8795 AACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATCAACCCACTCGTCACCCA 8854
Db |||||
QY 7820 ACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAACAGGAAGGC 7879
Db |||||
QY 8855 ACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAACAGGAAGGC 8914
Db |||||
QY 7880 AATAGTCCGCAAAAGGGAATAGGCGCACACGGAATGTGATCACTCATCTCTTCC 7939
Db |||||
QY 8915 AATAGTCCGCAAAAGGGAATAGGCGCGCACACGGAATGTGATCACTCATCTCTTCC 8974
Db |||||
QY 7940 TTTTTCATATTAATGAAGCAATTTATCAGGTTTATGTTCTCATGAGCGGATACATATTTG 7999
Db |||||
QY 8975 TTTTTCATATTAATGAAGCAATTTATCAGGTTTATGTTCTCATGAGCGGATACATATTTG 9034
Db |||||
QY 9000 AATGATTTAGAAATATAAAGTATAGGTTTCCGCGACATTTCCCGAAAGTGGCCAC 8059
Db |||||
QY 9035 AATGATTTAGAAATATAAAGTATAGGTTTCCGCGACATTTCCCGAAAGTGGCCAC 9094
Db |||||
QY 8060 CTGAGCTCTAAGAAACCAATTTATCATGACATTAACCTATAAAATAGGCGTATCACGA 8119
Db |||||
QY 9095 CTGAGCTCTAAGAAACCAATTTATCATGACATTAACCTATAAAATAGGCGTATCACGA 9154
Db |||||
QY 8120 GGCCTTTGCTTCAAGAAATTAATCTCATGTTGACAGTTTATCATGATTAAGCTGAC 8179
Db |||||
QY 9155 GGCCTTTGCTTCAAGAAATTAATCTCATGTTGACAGTTTATCATGATTAAGCTGAC 9214
Db |||||
QY 8180 TCATGTTGTTATGTAATAGACGCGATCGGAAACACTGAAATAAAGTATTAAT 8239
Db |||||
QY 9215 TCATGTTGTTATGTAATAGACGCGATCGGAAACACTGAAATAAAGTATTAAT 9274
Db |||||
QY 8240 CG 8241
Db |||||
QY 9275 CG 9276
Db |||||

RESULT 2

US-10-190-030B-13
; Sequence 13, Application US/10190030B
; Publication No. US20030134298A1
; GENERAL INFORMATION:
; APPLICANT: Madison, Edwin
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING A
; TITLE OF INVENTION: TRANSMEMBRANE SERINE PROTEASE 20, THE ENCODED POLYPEPTIDES AND
; FILE REFERENCE: 24745-1618
; CURRENT APPLICATION NUMBER: US/10/190,030B
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 9276
; TYPE: DNA
; ORGANISM: Pichia pastoris
US-10-190-030B-13

Query Match 79.6%; Score 6561.6; DB 15; Length 9276;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;

QY 1 AGATCTAATCATCCAAAGACGAAGGTTGAATGAACCTTTTGGCATCCGACATCCACAG 60
Db |||||
QY 1 AGATCTAATCATCCAAAGACGAAGGTTGAATGAACCTTTTGGCATCCGACATCCACAG 60
Db |||||
QY 61 GTCCATTCTCACAATAGTCCAAACGCAAGGAGGATACACTAGCAGCAGACCGT 120
Db |||||
QY 61 GTCCATTCTCACAATAGTCCAAACGCAAGGAGGATACACTAGCAGCAGACCGT 120
Db |||||
QY 121 TGCAAAACGAGGACCTCCACCTCTCTCTCAACACCCACCTTTTGGCATCGAAACCC 180
Db |||||
QY 121 TGCAAAACGAGGACCTCCACCTCTCTCTCAACACCCACCTTTTGGCATCGAAACCC 180
Db |||||
QY 181 AGCCAGTTATTTGGCTTGAATGGAGCTCGCTCATTCGAATTCCTTCTATTAGGCTACTA 240
Db |||||
QY 181 AGCCAGTTATTTGGCTTGAATGGAGCTCGCTCATTCGAATTCCTTCTATTAGGCTACTA 240
Db |||||
QY 241 ACACATGACCTTTATAGCCTCTCTATCTGGCCCTCTGGCAGGTTTCATGTTTCTTTA 300
Db |||||
QY 241 ACACATGACCTTTATAGCCTCTCTATCTGGCCCTCTGGCAGGTTTCATGTTTCTTTA 300
Db |||||
QY 301 TTTCCGAATGCAACGCTCGCATTTACCCGAAACATCACTCCAGATGAGGCTTTCTG 360
Db |||||
QY 361 AGTGTGGGTCATATAGTCTTTCATGTTTCCCAAAAGTCCCAAAACTGACAGTTTAAACGCT 420
Db |||||
QY 361 AGTGTGGGTCATATAGTCTTTCATGTTTCCCAAAAGTCCCAAAACTGACAGTTTAAACGCT 420
Db |||||
QY 421 GTCTTGGAACTTAATAGCAAAAGCGGTGATCTCATCCAAAGTGAACCTAAGTTTGGTTG 480
Db |||||
QY 481 TTGAAATGCTTAACCGCCAGTTGGTCAAAAGAAATCTCCAAAGTCCGACATACCGTTGT 540
Db |||||
QY 481 TTGAAATGCTTAACCGCCAGTTGGTCAAAAGAAATCTCCAAAGTCCGACATACCGTTGT 540
Db |||||
QY 541 CTGTTTGGTATTTGATGCAAAATGCTCAAAATTAATCTCATTAATGCTTAGCGAGTCT 600
Db |||||
QY 601 CTCTATCGCTTCTGAACCCCGGTGACCTGTGCGAAAGCAAAATGGGAAACACCCGCT 660
Db |||||
QY 601 CTCTATCGCTTCTGAACCCCGGTGACCTGTGCGAAAGCAAAATGGGAAACACCCGCT 660
Db |||||
QY 661 TTTTGGATGATTAATGCAATGCTTCCCAATGATGTTTCAAGATTTGGTGGGAATACT 720
Db |||||
QY 721 GCTGATAGCTTAAGCTTTCAATGATCAAAATTTAACTCTTCTTAACCTCTTCCAGCAAT 780
Db |||||
QY 721 GCTGATAGCTTAAGCTTTCAATGATCAAAATTTAACTCTTCTTAACCTCTTCCAGCAAT 780
Db |||||
QY 781 ATATAACAGAGGAGCTGCTCTGCTTAAACCTTTTATCATCATTTATAGCTT 840
Db |||||
QY 781 ATATAACAGAGGAGCTGCTCTGCTTAAACCTTTTATCATCATTTATAGCTT 840
Db |||||
QY 841 ACTTTCAATATTCGAGCTGTTTCCATTTGCAAGCTTTTGAATTTAAAGCTTTTAAACGA 900
Db |||||
QY 841 ACTTTCAATATTCGAGCTGTTTCCATTTGCAAGCTTTTGAATTTAAAGCTTTTAAACGA 900
Db |||||
QY 901 CAACTTGAGAGATCAAAACCACTAATTTATGAGGATCCAAACGATGAGATTTCT 960
Db |||||
QY 901 CAACTTGAGAGATCAAAACCACTAATTTATGAGGATCCAAACGATGAGATTTCT 960
Db |||||
QY 961 TCAATTTTACTGAGTTTATTTGCGAGCATCTCTCGCATTAGCTGCTCCAGTCAACT 1020
Db |||||
QY 961 TCAATTTTACTGAGTTTATTTGCGAGCATCTCTCGCATTAGCTGCTCCAGTCAACT 1020
Db |||||

Qy	1021	ACAA	CAGAGATGAAACGGCA	CAAAATTC	CGGCTGAAGCTGTCAT	TCGGTTACTCAGATT	TA 1081
Db	1021	ACAA	CAGAAGATGAAACGGCA	CAAAATTCGGGCTGAAGCTGTCAT	TCGGTTACTCAGATT	TA 1080	
Qy	1081	GAAGGGATTCGATGTTGCTGTTTTCGCA	TTTTTCCAAACGACCAAAATACGGGTTATTCG	1140			
Db	1081	GAAGGGATTCGATGTTGCTGTTTTCGCA	TTTTTCCAAACGACCAAAATACGGGTTATTCG	1140			
Qy	1141	TTTATAAATACTACTATTGCCAGCATTCGCTGCTA	AAAGAAGGGGTATCTCTCGAGAAA	1200			
Db	1141	TTTATAAATACTACTATTGCCAGCATTCGCTGCTA	AAAGAAGGGGTATCTCTCGAGAAA	1200			
Qy	1201	AGAGAGCTGAAGCCCAAGGCCAGGCCAGGCC	CCAGGCCCAAGGAAACAATGATCATG	1260			
Db	1201	AGAGAGCTGAAGCTTA-----	-----	1217			
Qy	1261	GCCCCCGGAGAGGATAAAATGTGGCTTC	CCCGGTGTCAACGCCACAGTCACCGGAG	1320			
Db	1218	-----	-----	1217			
Qy	1321	AGAGGTTGCTGTTTGTATGACAGTGT	CGGGGATTCGGTGTCTTCAACCCCATG	1380			
Db	1218	-----	-----	1217			
Qy	1381	ATCGAGAACACTCAAGAAGAAGATGTCC	CTTCTAACTAGTGGCGTAGAAATTC	1440			
Db	1218	-----	-----CGTAGAATTCCTAGGG	1234			
Qy	1441	CGGCGCGGAATTAATTCGCCCTTACAGATG	ACTGTTCCTCAGTTCAGTTGGGCACTTACG	1500			
Db	1235	CGGCGCGGAATTAATTCGCCCTTACAGATG	ACTGTTCCTCAGTTCAGTTGGGCACTTACG	1294			
Qy	1501	AGAAGACCGGTCCTTGCTAGATTCCTAAT	CAAGAGGATGTCAAGATGCCATTTGCTCGAG	1560			
Db	1295	AGAAGACCGGTCCTTGCTAGATTCCTAAT	CAAGAGGATGTCAAGATGCCATTTGCTCGAG	1354			
Qy	1561	ATGCAGGCTTCANTTTTGATCTTTTTTA	TTTGTAACTATATAGTATAGGATTTTTTTT	1620			
Db	1355	ATGCAGGCTTCANTTTTGATCTTTTTTA	TTTGTAACTATATAGTATAGGATTTTTTTT	1414			
Qy	1621	GTCAATTTGTTCTCTCGTACGAGCTGCT	CCTGATCAGCCATATCCGAGCTGATGA	1680			
Db	1415	GTCAATTTGTTCTCTCGTACGAGCTGCT	CCTGATCAGCCATATCCGAGCTGATGA	1474			
Qy	1681	TATCTTGCTAGGGGTTGGGAAAATCA	TTTCAGAGTTTGTATGTTTTTCTTTGGTATTTCCC	1740			
Db	1475	TATCTTGCTAGGGGTTGGGAAAATCA	TTTCAGAGTTTGTATGTTTTTCTTTGGTATTTCCC	1534			
Qy	1741	ACTCCTCTTCAGAGTACAGAAGATTA	AGTACAGAGTTTCGTTTGTGCAAGCTTATCGATA	1800			
Db	1535	ACTCCTCTTCAGAGTACAGAAGATTA	AGTACAGAGTTTCGTTTGTGCAAGCTTATCGATA	1594			
Qy	1801	GCITTTAATCGGGTAGTTTATCA	CAGTTTAAATTGCTAACGCAGTCAGGCACCGGTGATGA	1860			
Db	1595	GCITTTAATCGGGTAGTTTATCA	CAGTTTAAATTGCTAACGCAGTCAGGCACCGGTGATGA	1654			
Qy	1861	ATCTAACAATGGCGTCAATCGTCA	TCCTCGGCAACGTCACCTCGAATGCTTAGGCA	1920			
Db	1655	ATCTAACAATGGCGTCAATCGTCA	TCCTCGGCAACGTCACCTCGAATGCTTAGGCA	1714			
Qy	1921	CTTGTTATGCGGTACTGCGGGCCCT	TTTGGCGGATATCGTCAATTCGACAGCATCGC	1980			
Db	1715	CTTGTTATGCGGTACTGCGGGCCCT	TTTGGCGGATATCGTCAATTCGACAGCATCGC	1774			
Qy	1981	CAGTCATATGGCGTGTGCTAGCGCT	ATATGCGTTGATCAATTTCTATGCGCACCCCGT	2040			
Db	1775	CAGTCATATGGCGTGTGCTAGCGCT	ATATGCGTTGATCAATTTCTATGCGCACCCCGT	1834			
Qy	2041	TCTCGGAGCACTGTCGACACCGCT	TTTGGCGCGCCCGAGTCCGTGCTTCGCTACTTGG	2100			
Db	1835	TCTCGGAGCACTGTCGACACCGCT	TTTGGCGCGCCCGAGTCCGTGCTTCGCTACTTGG	1894			
Qy	2101	AGCCACTATCGACTACCGCATCAT	TGGCGGACCAACACCCGCTCGTGTGATCTATCGAATCTA	2160			

1895	Db		AGCCACTATTCGACTATGCGGATCATGGCGACCAACCGCTCTCTGTGATCTTATCGAATCTA	1954
2161	Qy		AATGTAAGTTAAATCTCTAAATAAATTAATAATAGTCCAGGTTTCTCCATACGAACTTAA	2220
1955	Db		AATGTAAGTTAAATCTCTAAATAAATTAATAATAGTCCAGGTTTCTCCATACGAACTTAA	2014
2221	Qy		CAGCATTCGGGTGAGCATCTPAGACCTTCAACAGCAGCCAGATCCATCACTGCTTGGCCAA	2280
2015	Db		CAGCATTCGGGTGAGCATCTPAGACCTTCAACAGCAGCCAGATCCATCACTGCTTGGCCAA	2074
2281	Qy		TATGTTTTCAGTCCCTCAGGAGTTACGTCCTGTCGAAGTGATGAATCTCTGGAAGGTTGCAG	2340
2075	Db		TATGTTTTCAGTCCCTCAGGAGTTACGTCCTGTCGAAGTGATGAATCTCTGGAAGGTTGCAG	2134
2341	Qy		TGTTAACTCCGCTGTATTTGACGGGCATATCCGTACGTTGGCAAAGTGTGGTACCGG	2400
2135	Db		TGTTAACTCCGCTGTATTTGACGGGCATATCCGTACGTTGGCAAAGTGTGGTACCGG	2194
2401	Qy		AGGAGTAATCTCCACAACCTCTCTGGAGAGTAGGCCACCAACAACACAGATCCAGCGTGT	2460
2195	Db		AGGAGTAATCTCCACAACCTCTCTGGAGAGTAGGCCACCAACAACACAGATCCAGCGTGT	2354
2461	Qy		GTACTTGATCAACANTAAGAAAGCAATTCCTCGATTTGCAAGTCAAGTGTTCAGGAGCGT	2520
2255	Db		GTACTTGATCAACANTAAGAAAGCAATTCCTCGATTTGCAAGTCAAGTGTTCAGGAGCGT	2314
2521	Qy		ACTGATTGGACATTTCCAAAGCCTGCTGTGAGGTTGCAACCGATAGGTTGTAGAGTGTC	2580
2315	Db		ACTGATTGGACATTTCCAAAGCCTGCTGTGAGGTTGCAACCGATAGGTTGTAGAGTGTC	2374
2581	Qy		CAATACACTTGGCTGCAATTTCAACCTTGGCAACTGCAAGCTTGGTGTGTAACAGCAT	2640
2375	Db		CAATACACTTGGCTGCAATTTCAACCTTGGCAACTGCAAGCTTGGTGTGTAACAGCAT	2434
2641	Qy		CTTCAATTCGGCAAGCTCTCTGTGCATATCGACAGCCACAGAATCACTGCGGAAT	2700
2435	Db		CTTCAATTCGGCAAGCTCTCTGTGCATATCGACAGCCACAGAATCACTGCGGAAT	2494
2701	Qy		CAATACACTTTCAGCTTTGAGACAGAGGTCCTGAGCAACGAATCTGGATCAGCGTATT	2760
2495	Db		CAATACACTTTCAGCTTTGAGACAGAGGTCCTGAGCAACGAATCTGGATCAGCGTATT	2554
2761	Qy		TATCAGCAATTAACCTAGAACTTCAGAAAGGCCACAGGCGATGTCAATACTACAGAGGCTG	2820
2555	Db		TATCAGCAATTAACCTAGAACTTCAGAAAGGCCACAGGCGATGTCAATACTACAGAGGCTG	2614
2821	Qy		ATGTGTCATTTGAAACCATCATCTTGGCAGCAGTACGAAGTCTTCTGGACCAATA	2880
2615	Db		ATGTGTCATTTGAAACCATCATCTTGGCAGCAGTACGAAGTCTTCTGGACCAATA	2674
2881	Qy		TTTTTGCACACTTAGAAACAGTTTCTGTTCGGTAAAGCCATAGCAGCTACTGCTCTGGCGC	2940
2675	Db		TTTTTGCACACTTAGAAACAGTTTCTGTTCGGTAAAGCCATAGCAGCTACTGCTCTGGCGC	2734
2941	Qy		CTCCTGTTAGCAGATACACTTAGACACCACTTGTGGGCAACGTFAGATGACTTCTGGGG	3000
2735	Db		CTCCTGTTAGCAGATACACTTAGACACCACTTGTGGGCAACGTFAGATGACTTCTGGGG	2794
3001	Qy		TAAGGTTACCATCTCTTAGGTGGAGATGCAAAAACAAATTTCTTTGCAACCAAGCAACTT	3060
2795	Db		TAAGGTTACCATCTCTTAGGTGGAGATGCAAAAACAAATTTCTTTGCAACCAAGCAACTT	2854
3061	Qy		TGCGAGAAACACCCAGCATCAGGGAAGTGGAAAGCGAGAAATTCGGTTTCCACAGGAATAT	3120
2855	Db		TGCGAGAAACACCCAGCATCAGGGAAGTGGAAAGCGAGAAATTCGGTTTCCACAGGAATAT	2914
3121	Qy		AGAGGCCAATTTCTCAATAGGCTTTGCAAAACGAGAGCAGACTACACGAGGCGAAGTCT	3180
2915	Db		AGAGGCCAATTTCTCAATAGGCTTTGCAAAACGAGAGCAGACTACACGAGGCGAAGTCT	2974
3181	Qy		CAACTTGCACACGCTCCGTTAGTTGAGCTTCATGGAATTTCTGACGTTATCTATAGAGA	3240

Db 2975 CAACCTTCCAAAGCTCTCGTTAGTTGAGCTTCATGGAATTTCTGAGCTTATCTATAGAGA 3034
Qy 3241 GATCAATGGCTCTCTTAACGTTATCTGGCAATTTGCAATAGTTCTCTCTGGGAAAGGAGCTT 3300
Db 3035 GATCAATGGCTCTCTTAACGTTATCTGGCAATTTGCAATAGTTCTCTCTGGGAAAGGAGCTT 3094
Qy 3301 CTAACACAGGTGCTCTCAAGCGACTCCATCAACCTTGGCAGTTAGTTCTTAAAGGGCTT 3360
Db 3095 CTAACACAGGTGCTCTCAAGCGACTCCATCAACCTTGGCAGTTAGTTCTTAAAGGGCTT 3154
Qy 3361 TGTCAACATTTTGAACGAACATTTGCGACAAATTTGGTTGACTAATTCATAATCTGTTCGG 3420
Db 3155 TGTCAACATTTTGAACGAACATTTGCGACAAATTTGGTTGACTAATTCATAATCTGTTCGG 3214
Qy 3421 TTTTCTCGATAGGACGAGGAGGCACTTCAATTTCTGTGAGGAGGCTTGAACCTT 3480
Db 3215 TTTTCTCGATAGGACGAGGAGGCACTTCAATTTCTGTGAGGAGGCTTGAACCTT 3274
Qy 3481 CAATTTTGCACAAATCAATACGACTTTCAGAGGAGCTTCTTTAGGTTTGGATTCTTCTT 3540
Db 3275 CAATTTTGCACAAATCAATACGACTTTCAGAGGAGCTTCTTTAGGTTTGGATTCTTCTT 3334
Qy 3541 TAGGTTGTTCTTGGTGATCTCTGGCTTGGGATCTCTCTTCTCTTCTAGTGACCTTTAGGG 3600
Db 3335 TAGGTTGTTCTTGGTGATCTCTGGCTTGGGATCTCTCTTCTCTTCTAGTGACCTTTAGGG 3394
Qy 3601 ACTTCATATCCAGGTTTCTCTCCACCTCTGCCAACGTCACACGCTACTTGGCACATCTAA 3660
Db 3395 ACTTCATATCCAGGTTTCTCTCCACCTCTGCCAACGTCACACGCTACTTGGCACATCTAA 3454
Qy 3661 CTAATGCAAAAATAAATAAGTCAGCACATCCCAAGGCTATATCTTCTTGGATTAGCTT 3720
Db 3455 CTAATGCAAAAATAAATAAGTCAGCACATCCCAAGGCTATATCTTCTTGGATTAGCTT 3514
Qy 3721 CTGCAAGTTCATCAGCTTCTCTCCCTAATTTTAGGGTTCAACAAATCTCTGCTCAAAATA 3780
Db 3515 CTGCAAGTTCATCAGCTTCTCTCCCTAATTTTAGGGTTCAACAAATCTCTGCTCAAAATA 3574
Qy 3781 ACCGTTTGGTATAAGAACCTTCTGGAGCATGCTCTTACGATCCCAACAGGTGGCTTCCA 3840
Db 3575 ACCGTTTGGTATAAGAACCTTCTGGAGCATGCTCTTACGATCCCAACAGGTGGCTTCCA 3634
Qy 3841 TGGCTCTAAGACCTTTTGATTTGGCCAAAACAGGAAGTGGCTTCAAGTCACAGAACCAA 3900
Db 3635 TGGCTCTAAGACCTTTTGATTTGGCCAAAACAGGAAGTGGCTTCAAGTCACAGAACCAA 3694
Qy 3901 CACCTGTTTGTTCACCAACAAATTTCAAGCAGTCTCCATCACAATCCAATTCGATACCCA 3960
Db 3695 CACCTGTTTGTTCACCAACAAATTTCAAGCAGTCTCCATCACAATCCAATTCGATACCCA 3754
Qy 3961 GCACTTTTGTGATGCTCCAGATGATAGCACTTTATACCAACAAACCTGTGACGAGATT 4020
Db 3755 GCACTTTTGTGATGCTCCAGATGATAGCACTTTATACCAACAAACCTGTGACGAGATT 3814
Qy 4021 GGTAGACTCCAGTTTGTGCTTATAGCTCTCGGAATAGACTTTTGGACGAGTACACCA 4080
Db 3815 GGTAGACTCCAGTTTGTGCTTATAGCTCTCGGAATAGACTTTTGGACGAGTACACCA 3874
Qy 4081 GGCCCAACGAGTAATFAGAAGAGTCAGCCACAAAGTAGTGAATFAGACCATCGGGGGCTT 4140
Db 3875 GGCCCAACGAGTAATFAGAAGAGTCAGCCACAAAGTAGTGAATFAGACCATCGGGGGCTT 3934
Qy 4141 CAGTAGTCAAGAGCCCAACAAATTTCACTGACAGGGAACTTTTTCACATCTTTCAGAAA 4200
Db 3935 CAGTAGTCAAGAGCCCAACAAATTTCACTGACAGGGAACTTTTTCACATCTTTCAGAAA 3994
Qy 4201 GTTTCGTATTCAGTAGTCAATTTGCCGAGCATCAATAATGGGATTATACCAGAGCAACAG 4260
Db 3995 GTTTCGTATTCAGTAGTCAATTTGCCGAGCATCAATAATGGGATTATACCAGAGCAACAG 4054
Qy 4261 TGGAGTCACTACCTACCACTTTGGGCTCTCAGAAAAGCATAAACAGTTCTTACTACCGC 4320
Db 4055 TGGAGTCACTACCTACCACTTTGGGCTCTCAGAAAAGCATAAACAGTTCTTACTACCGC 4114

Qy 4321 CATTAGTGAACCTTTTCAAAATCGCCAGTCGAGAGAGAAAAGGCACAGCGATACTAGCAT 4380
Db 4115 CATTAGTGAACCTTTTCAAAATCGCCAGTCGAGAGAGAAAAGGCACAGCGATACTAGCAT 4174
Qy 4381 TAGCGGGCAAGGATGCAACTTTTATCAACCGAGGTCCTATAGATAAACCTTAGCCGCTGGGA 4440
Db 4175 TAGCGGGCAAGGATGCAACTTTTATCAACCGAGGTCCTATAGATAAACCTTAGCCGCTGGGA 4234
Qy 4441 TCATCTCTTTGACAACTCTTTCTGCCAAATCTAGTCCAAAATCACCTTCAATTTGATACCAT 4500
Db 4235 TCATCTCTTTGACAACTCTTTCTGCCAAATCTAGTCCAAAATCACCTTCAATTTGATACCAT 4294
Qy 4501 TATTGTACAACCTTGAGCAAGTTGTGATCGATCGACTCCTCAAAATTTGGTCTCTGTAAACGATG 4560
Db 4295 TATTGTACAACCTTGAGCAAGTTGTGATCGATCGACTCCTCAAAATTTGGTCTCTGTAAACGATG 4354
Qy 4561 ACTCAACTTGCACATTTAACTTTGAGCTCAGTCGATTTGAGTGAACCTTGCATCAGGTTGTGCA 4620
Db 4355 ACTCAACTTGCACATTTAACTTTGAGCTCAGTCGATTTGAGTGAACCTTGCATCAGGTTGTGCA 4414
Qy 4621 GCTGCTCAGCAGCATAGGGAACACGCGCTTTTCTACCAAACTCAAGGAATTTATCAAACT 4680
Db 4415 GCTGCTCAGCAGCATAGGGAACACGCGCTTTTCTACCAAACTCAAGGAATTTATCAAACT 4474
Qy 4681 CTGCAACACTTTGGCGTATGCGAGTAGCAAGGAATGTCTACTTTGAAGTCGACAGTGAG 4740
Db 4475 CTGCAACACTTTGGCGTATGCGAGTAGCAAGGAATGTCTACTTTGAAGTCGACAGTGAG 4534
Qy 4741 TGTAAGTTGAGAAATCTGAAGCGCTATTTTATTTATCAGTGAAGTCAGTCATCAGGAGA 4800
Db 4535 TGTAAGTTGAGAAATCTGAAGCGCTATTTTATTTATCAGTGAAGTCAGTCATCAGGAGA 4594
Qy 4801 TCTCTACCGCGGACGATCTGTGCGGA----- 4828
Db 4595 TCTCTACCGCGGACGATCTGTGCGCGCACTCGTGGCGGCGGGGGGGGGGGGGCTGTGC 4654
Qy 4829 ----- 4828
Db 4655 CTCTGGAAGAGGTGTTGCTGACTCATACAGGCTTGATCGCCCATCATCCAGCCAGA 4714
Qy 4829 ----- 4828
Db 4715 AAGTGAAGGAGCCAGGTTGATGAGAGCTTTGTTGTTAGGTGGACCGAGTTGGTATTGA 4774
Qy 4829 ----- 4828
Db 4775 ACTTTTGTCTTCCACGGAACGGTCTGCTGTTGTCGGGAAGATGCGTGATCTGATCCTTCA 4834
Qy 4829 ----- 4828
Db 4835 ACTCAGCAAAAGTTGCAATTTATTAACAAAGCGCGCTCCGCTCAAGTCAGGTAATGCT 4894
Qy 4829 ----- 4828
Db 4895 CTGCCAGTTTACAAACCAATTAACCAATTTGATTAGAAAAAATCATCGAGCATCAAAATG 4954
Qy 4829 ----- 4828
Db 4955 AAATGCAATTTTATTCATATCAGGATTATCAATACCATATTTTTCGAAAAAGCGGTTTCTG 5014
Qy 4829 ----- 4828
Db 5015 TAATGAAGGAGAAAACTCACCAGGAGTTCATAGTAGGCAAGATCTCTGGTATCGGTC 5074
Qy 4829 ----- 4828
Db 5075 TGGGATTCGACTCGTCCCAACATCAATACAACTTATTATTTCCCTCGTCAAAAAATAAG 5134
Qy 4829 ----- 4828
Db 5135 GTTATCAAGTGAGAAATCACCATTAGTACGAGTGAATCCGGTGAAGTGGCAAAAGCTT 5194

7355 GGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC 7414
 6380 CGACAGACTATAAGATACAGCGCTTCCCGCTGGAAGCTCCCTCGTGGCTCTCCCTG 6439
 7415 CGACAGACTATAAGATACAGCGCTTCCCGCTGGAAGCTCCCTCGTGGCTCTCCCTG 7474
 6440 TTCCGACCTCCGCTTACCGGATACCTGTCGCTTTCCTCCCTGGAAGCGTGGCGC 6499
 7475 TTCCGACCTCCGCTTACCGGATACCTGTCGCTTTCCTCCCTGGAAGCGTGGCGC 7534
 6500 TTCTCTAGTCTACGCTGTAGTATCTCAGTTTCGCTGAGTTCGCTCCCTCAAGCTGG 6559
 7535 TTCTCTAGTCTACGCTGTAGTATCTCAGTTTCGCTGAGTTCGCTCCCTCAAGCTGG 7594
 6560 GCTGTGTGACGAACCCCGCTTACCGGATACCTGTCGCTTTCCTCCCTGGAAGCGTGGCGC 6619
 7595 GCTGTGTGACGAACCCCGCTTACCGGATACCTGTCGCTTTCCTCCCTGGAAGCGTGGCGC 7654
 6620 TTGAGTCCAAACCCCGCTTACCGGATACCTGTCGCTTTCCTCCCTGGAAGCGTGGCGC 6679
 7655 TTGAGTCCAAACCCCGCTTACCGGATACCTGTCGCTTTCCTCCCTGGAAGCGTGGCGC 7714
 6680 TTACGACGAGGAGTATGAGGCGCTGCTACAGATTCCTTGAAGTGGTGGCTTAACTACG 6739
 7715 TTACGACGAGGAGTATGAGGCGCTGCTACAGATTCCTTGAAGTGGTGGCTTAACTACG 7774
 6740 GCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA 6799
 7775 GCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA 7834
 6800 AAAGAGTTGGTGTAGTCTTGTATCCCGGCAACCAACACCGCTGTAGCGGTGTTTTTTT 6859
 7835 AAAGAGTTGGTGTAGTCTTGTATCCCGGCAACCAACACCGCTGTAGCGGTGTTTTTTT 7894
 6860 TTTCGAAGCAGCAGATTCAGCGCAGAAAAAAGGATCTCAAGAGATTCCTTTGATCTTTT 6919
 7895 TTTCGAAGCAGCAGATTCAGCGCAGAAAAAAGGATCTCAAGAGATTCCTTTGATCTTTT 7954
 6920 CTACGGGTCTGAGCTCAGTGGAGCAAGAACTCAGCTTAAGGATTTTGGTCTATGAGAT 6979
 7955 CTACGGGTCTGAGCTCAGTGGAGCAAGAACTCAGCTTAAGGATTTTGGTCTATGAGAT 8014
 6980 TATCAAAAAGGATCTTACCTAGATCTCTTTTAAATTAATAATGAAGTTTAAATCAATCT 7039
 8015 TATCAAAAAGGATCTTACCTAGATCTCTTTTAAATTAATAATGAAGTTTAAATCAATCT 8074
 7040 AAAGPATATAGTAATCTTGGTCTGACGATTAACCAATGTTAATCAGTGGCCACTTA 7099
 8075 AAAGPATATAGTAATCTTGGTCTGACGATTAACCAATGTTAATCAGTGGCCACTTA 8134
 7100 TCTACGAGATCTGCTATTTTCTGCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7159
 8135 TCTACGAGATCTGCTATTTTCTGCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8194
 7160 CTACGATACGGAGGCTTACCATCTGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7219
 8195 CTACGATACGGAGGCTTACCATCTGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8254
 7220 GCTACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCGGAGCGGAGCGGAGGAGGAGGAG 7279
 8255 GCTACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCGGAGCGGAGCGGAGGAGGAGGAG 8314
 7280 GTGGTCTCGCACTTTATCCGCTCCATCCAGTCTTATTAATTTGTCGGGGAAGCTAGAG 7339
 8315 GTGGTCTCGCACTTTATCCGCTCCATCCAGTCTTATTAATTTGTCGGGGAAGCTAGAG 8374
 7340 TAAGTAGTTCGCACTTAATAGTTTGGCAAGGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7399
 8375 TAAGTAGTTCGCACTTAATAGTTTGGCAAGGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8434
 7400 TGTACGCTCGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 7459
 8435 TGTACGCTCGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 8494

7460 TTACATGATCCCATCTTTGTGCAAAAAAGCGGTAGTCTCTTCCGCTCCCTCCGATCGTTG 7519
 8495 TTACATGATCCCATCTTTGTGCAAAAAAGCGGTAGTCTCTTCCGCTCCCTCCGATCGTTG 8554
 7520 TCAGAAGTAAAGTTTCGCGCAGTGTATCACTCATCTGTTATGCGAGCAGCTGCAATTTCTC 7579
 8555 TCAGAAGTAAAGTTTCGCGCAGTGTATCACTCATCTGTTATGCGAGCAGCTGCAATTTCTC 8614
 7580 TTACTGTCAAGCCATCCGTAAGATGCTTTTCTGTGACTGTGAGTACTCAACCAAGTCAT 7639
 8615 TTACTGTCAAGCCATCCGTAAGATGCTTTTCTGTGACTGTGAGTACTCAACCAAGTCAT 8674
 7640 TCTGAGTATAGTATGCGGCGACCGAGTTGCTCTTCCCGCGCTCAACACGGGATATA 7699
 8675 TCTGAGTATAGTATGCGGCGACCGAGTTGCTCTTCCCGCGCTCAACACGGGATATA 8734
 7700 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAACAAAGTTCTTCGGGGCGAA 7759
 8735 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAACAAAGTTCTTCGGGGCGAA 8794
 7760 AACTCTCAAGATCTTACCGCTGTGAGTCCAGTTCGATGTAAACCCACTCGTGCAACCCA 7819
 8795 AACTCTCAAGATCTTACCGCTGTGAGTCCAGTTCGATGTAAACCCACTCGTGCAACCCA 8854
 7820 ACTGATCTTACGATCTTTTACTTTTACCCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGC 7879
 8855 ACTGATCTTACGATCTTTTACTTTTACCCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGC 8914
 7880 AAATGCGCGCAAAAAGGGAATTAAGGCGGACACCGAAATGTTGAATACTCATCTCTTCC 7939
 8915 AAATGCGCGCAAAAAGGGAATTAAGGCGGACACCGAAATGTTGAATACTCATCTCTTCC 8974
 7940 TTTTTCATATTTTGAAGCATTTTATCAGGCTTATTTGCTCATGAGCGGATACATATTTG 7999
 8975 TTTTTCATATTTTGAAGCATTTTATCAGGCTTATTTGCTCATGAGCGGATACATATTTG 9034
 8000 AATGATTTAGAAAAATAAACAATAGGCTTCCGCGCAGATTTCCCGGAAAAAGTCCAC 8059
 9035 AATGATTTAGAAAAATAAACAATAGGCTTCCGCGCAGATTTCCCGGAAAAAGTCCAC 9094
 8060 CTGAGCTCTAAGAACCAATTTATCATGATTAACCTATAAATAAGGCGTATCAGA 8119
 9095 CTGAGCTCTAAGAACCAATTTATCATGATTAACCTATAAATAAGGCGTATCAGA 9154
 8120 GGCCCTTTCTGCTTCAAGATTAATTTCTGTTGAGCTTATCATCTGATTAAGCTGAC 8179
 9155 GGCCCTTTCTGCTTCAAGATTAATTTCTGTTGAGCTTATCATCTGATTAAGCTGAC 9214
 8180 TCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 8239
 9215 TCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 9274
 8240 CG 8241
 9275 CG 9276

RESULT 3
 US-10-302-840A-13
 : Sequence 13, Application US/10302840A
 : Publication No. US20030134794A1
 : GENERAL INFORMATION:
 : APPLICANT: Madison, Edwin L.
 : APPLICANT: Ong, Edgar O.
 : TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING SERINE PROTEASE CVSP17, THE ENCODED
 : TITLE OF INVENTION: POLYPEPTIDES AND METHODS BASED THEREON
 : FILE REFERENCE: 24745-1622
 : CURRENT APPLICATION NUMBER: US/10/302,840A
 : CURRENT FILING DATE: 2003-01-24
 : PRIOR APPLICATION NUMBER: 60/332,015
 : PRIOR FILING DATE: 2001-11-20
 : NUMBER OF SEQ ID NOS: 18

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 9276
; TYPE: DNA
; ORGANISM: Pichia pastoris
US-10-302-840A-13

Query Match      79.6%; Score 6561.6; DB 15; Length 9276;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;

QY 1 AGATCTAACATCCAAAGAGCAAGAGGTTGAATGAACCTTTTGGCATCCGACATCCACAG 60
Db 1 AGATCTAACATCCAAAGAGCAAGAGGTTGAATGAACCTTTTGGCATCCGACATCCACAG 60
QY 61 GTCCATTCTCACATAGTSCAAAACGCAACAGAGGGGATACACTAGCAGACAGCGT 120
Db 61 GTCCATTCTCACATAGTSCAAAACGCAACAGAGGGGATACACTAGCAGACAGCGT 120
QY 121 TGCAAACGAGAGCTCCACTCCTCTCTCTCAACCCACCTTTTGGCATCCGAAACAC 180
Db 121 TGCAAACGAGAGCTCCACTCCTCTCTCTCAACCCACCTTTTGGCATCCGAAACAC 180
QY 181 AGCCCAAGTTATGGGCTTGAATGGAGCTCGCTCATTCCTCAATTCCTCTATTAGGCTACTA 240
Db 181 AGCCCAAGTTATGGGCTTGAATGGAGCTCGCTCATTCCTCAATTCCTCTATTAGGCTACTA 240
QY 241 ACACATGACTTTATAGCCTGTCTATCTGCGCCCGCTGCGAGGTTCAITGTTGTTA 300
Db 241 ACACATGACTTTATAGCCTGTCTATCTGCGCCCGCTGCGAGGTTCAITGTTGTTA 300
QY 301 TTTCCGAATCAACAGCTCGGCTTACACCCGCAACATCACTCCAGATGAGGGCTTTCTG 360
Db 301 TTTCCGAATCAACAGCTCGGCTTACACCCGCAACATCACTCCAGATGAGGGCTTTCTG 360
QY 361 AGTGCGGGTCAATAGTTTCATGTTCCCAAAATGGGCCAAAACCTGACAGTTAAACGCT 420
Db 361 AGTGCGGGTCAATAGTTTCATGTTCCCAAAATGGGCCAAAACCTGACAGTTAAACGCT 420
QY 421 GTCCTTGAACCTATATGACAAAGCGTGATCTCATCCAGATCAACTAAGTTGGTTCG 480
Db 421 GTCCTTGAACCTATATGACAAAGCGTGATCTCATCCAGATCAACTAAGTTGGTTCG 480
QY 481 TTGAAATGCTTAACGGCCAGTTGGTCAAAAAGAAACTTCCAAAAGTCCGCAATACCGTTGT 540
Db 481 TTGAAATGCTTAACGGCCAGTTGGTCAAAAAGAAACTTCCAAAAGTCCGCAATACCGTTGT 540
QY 541 CTGTTTGGTATGATGACCAATGCTCAAAAATTAATCTCAATTAATGCTTAGCGAGTCT 600
Db 541 CTGTTTGGTATGATGACCAATGCTCAAAAATTAATCTCAATTAATGCTTAGCGAGTCT 600
QY 601 CTCTATCGCTTCTGAAACCCCGTGACCTGTGCGGAAACGCAAAATGGGAAACACCCGCT 660
Db 601 CTCTATCGCTTCTGAAACCCCGTGACCTGTGCGGAAACGCAAAATGGGAAACACCCGCT 660
QY 661 TTTTGGATGATATGCAATGTCTCCACATTTGTATGCTTCCAAAGATTTCTGGTGGAACT 720
Db 661 TTTTGGATGATATGCAATGTCTCCACATTTGTATGCTTCCAAAGATTTCTGGTGGAACT 720
QY 721 GCTGATGCTTAAACGTTCAATGATCAAAATTTAACTGTTCTAACCCCTACTTGCAGCAAT 780
Db 721 GCTGATGCTTAAACGTTCAATGATCAAAATTTAACTGTTCTAACCCCTACTTGCAGCAAT 780
QY 781 ATATAACAGAGGAGCTGCCCTGTCTTAAACCTTTTTTTTATCATCATTAATAGCTT 840
Db 781 ATATAACAGAGGAGCTGCCCTGTCTTAAACCTTTTTTTTATCATCATTAATAGCTT 840
QY 841 ACTTTCATATGCGACTGGTTCGAATGACAGCTTTTGATTTTAAAGACTTTTAACGA 900
Db 841 ACTTTCATATGCGACTGGTTCGAATGACAGCTTTTGATTTTAAAGACTTTTAACGA 900
QY 901 CAACCTTGAGAGATCAAAAACCACTAATTTATCGAAGATCCAAACGATGAGATTTTCT 960
Db 901 CAACCTTGAGAGATCAAAAACCACTAATTTATCGAAGATCCAAACGATGAGATTTTCT 960

901 CAACTTGAGAGATCAAAAACCACTAATTTATCGAAGATCCAAACGATGAGATTTCT 960
961 TCAATTTTACTGCAAGTTTATTCGAGCATCTCCGCAATGAGCTGCTCCAGTCAACT 1020
961 TCAATTTTACTGCAAGTTTATTCGAGCATCTCCGCAATGAGCTGCTCCAGTCAACT 1020
1021 ACACAGAGATGAAACGCAACATTCGGCTGAGCTGCTCATCGTGTACTCAGATTTA 1080
1021 ACACAGAGATGAAACGCAACATTCGGCTGAGCTGCTCATCGTGTACTCAGATTTA 1080
1081 GAAGGGATTTTCGATGTTGCTGTTTGGCATTTTCCACAGCAAAATAACGGGTTATG 1140
1081 GAAGGGATTTTCGATGTTGCTGTTTGGCATTTTCCACAGCAAAATAACGGGTTATG 1140
1141 TTTATAATACACTATTGCGCAGCATTCGCTGCTAAGAGAGGGGTATCTCTCAGAAA 1200
1141 TTTATAATACACTATTGCGCAGCATTCGCTGCTAAGAGAGGGGTATCTCTCAGAAA 1200
1201 AGAGAGGCTGAAAGCCAGGCCAGGCCAGGCCAGGCCAGGAAAGAAACATGTATCATG 1260
1201 AGAGAGGCTGAAAGCTTAA----- 1217
1261 GCGCCCGGAGAGAGATAAATTTGCTGCTTCCCGGCTGTCCCGCCAGCAGTGCAGGAG 1320
1218 ----- 1217
1321 AGAGGTTGCTGTTTGTATGACAGCTGTCGGGGATTCGGGTGCTTCCACCCCATGGCC 1380
1218 ----- 1217
1381 ATCGAACACTCAAGAAGAAGATGTCCTTCTAACTAGTGGCGTAGAATTTCCCTAGGG 1440
1218 -----CGTAGAATTTCCCTAGGG 1234
1441 CGCCCGGAAATTAATTCGCTTAGACATGACTGTTCTCAGTTCAGTTGAGTGGCAGCTTACG 1500
1235 CGCCCGGAAATTAATTCGCTTAGACATGACTGTTCTCAGTTCAGTTGAGTGGCAGCTTACG 1294
1501 AGAAGACCCGCTCTGCTAGATTTCTAATCAAGAGGATGTGAGATGCAATTTGCCTGAGAG 1560
1295 AGBAGACCCGCTCTGCTAGATTTCTAATCAAGAGGATGTGAGATGCAATTTGCCTGAGAG 1354
1561 ATCGAGCTTCATTTTGTATGACTTTTATTTTATTTGTAACCTATATAGTATAGGATTTTTT 1620
1355 ATCGAGCTTCATTTTGTATGACTTTTATTTTATTTGTAACCTATATAGTATAGGATTTTTT 1414
1621 GTCATTTTCTTCTCTGCTAGAGCTGCTCTGATCAGCCTATCTCGAGCTGATGAA 1680
1415 GTCATTTTGTCTCTCTGCTAGAGCTGCTCTGATCAGCCTATCTCGAGCTGATGAA 1474
1681 TATCTTTGCTAGGGGTTTGGGAAATCATTCGAGTTGATGTTTCTTTGTTATTTTCCC 1740
1475 TATCTTTGCTAGGGGTTTGGGAAATCATTCGAGTTGATGTTTCTTTGTTATTTCCC 1534
1741 ACTCCTCTTCAGAGTACAGAAGATTAAGTGAGAAGTTGCTTTGTCAGCTTATCGATAA 1800
1535 ACTCCTCTTCAGAGTACAGAAGATTAAGTGAGAAGTTGCTTTGTCAGCTTATCGATAA 1594
1801 GCTTTAATCGGCTAGTTTATCAGTTAAATTTGCTAAACAGCTCAGGCAACCGTGTATGAA 1860
1595 GCTTTAATCGGCTAGTTTATCAGTTAAATTTGCTAAACAGCTCAGGCAACCGTGTATGAA 1654
1861 ATCTAACAATGGCTCATCTGCTCGCAGCCGCTGATGCTGCTGAGGATGCTGCTAGGATAG 1920
1655 ATCTAACAATGGCTCATCTGCTCGCAGCCGCTGATGCTGCTGAGTGTGCTAGGATAGG 1714
1921 CTTGGTTATGCGGCTACTGCGGGCTCTTGGGGATATCGTCCATTTCCGACAGCATCGC 1980
1715 CTTGGTTATGCGGCTACTGCGGGCTCTTGGGGATATCGTCCATTTCCGACAGCATCGC 1774
1981 CAGTCACTATGGCGTCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCGCT 2040
1775 CAGTCACTATGGCGTCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCGCT 1834
```

QY 2041 TCTCGGAGCACTGTCCGACCGCTTTGGCCGCGCCGAGTCTCTGCTCGCTTCGCTACTTGG 2100
Db 1835 TCTCGGAGCACTGTCCGACCGCTTTGGCCGCGCCGAGTCTCTGCTCGCTACTTGG 1894
QY 2101 AGCCACTATCGACTACCGCATCATGGCGACCAACACCGCTCTCTGCGATCTATCGAATCTA 2160
Db 1895 AGCCACTATCGACTACCGCATCATGGCGACCAACACCGCTCTCTGCGATCTATCGAATCTA 1954
QY 2161 AATGTAAGTTAAATCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2220
Db 1955 AATGTAAGTTAAATCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2014
QY 2221 CAGCAATTCGGTGAGCATCTAGACCTTCAACAGCAGCCAGATCCATCACTGCTTGCCCAA 2280
Db 2015 CAGCAATTCGGTGAGCATCTAGACCTTCAACAGCAGCCAGATCCATCACTGCTTGCCCAA 2074
QY 2281 TATGTTTCAGTCCCTCAGGAGTTACGCTTGTGAAGTGATGAACCTTCTGGAAGGTTGCAG 2340
Db 2075 TATGTTTCAGTCCCTCAGGAGTTACGCTTGTGAAGTGATGAACCTTCTGGAAGGTTGCAG 2134
QY 2341 TGTTAACTCCGCTGTATGAGCGGCATATCGGTACGTTGGCAAGTGTGGTACCGG 2400
Db 2135 TGTTAACTCCGCTGTATGAGCGGCATATCGGTACGTTGGCAAGTGTGGTACCGG 2194
QY 2401 AGGAGTAATCTCCAACTCTCTCGGAGTAGGCAACCAACACAGATCCAGCGTGT 2460
Db 2195 AGGAGTAATCTCCAACTCTCTCGGAGTAGGCAACCAACACAGATCCAGCGTGT 2254
QY 2461 GTACTGATCAACATAGAGAGCATTCGATTTGCAGGATCAAGTGTTCAGGAGCGT 2520
Db 2255 GTACTGATCAACATAGAGAGCATTCGATTTGCAGGATCAAGTGTTCAGGAGCGT 2314
QY 2521 ACTGATTTGAGCATTTCCAAAGCTCTGCTAGGTTGCAACCGATAGGTTGTAGAGTGTG 2580
Db 2315 ACTGATTTGAGCATTTCCAAAGCTCTGCTAGGTTGCAACCGATAGGTTGTAGAGTGTG 2374
QY 2581 CAATACTTGGTACCAATTTCAACCTCTGGCAACTGCAAGCTTGGTGTGTAACAGCAT 2640
Db 2375 CAATACTTGGTACCAATTTCAACCTCTGGCAACTGCAAGCTTGGTGTGTAACAGCAT 2434
QY 2641 CTTCAATTCGGCAGCTCTCTGCTGTCATATCGACAGCAACAGAAATCACTGCGCAAT 2700
Db 2435 CTTCAATTCGGCAGCTCTCTGCTGTCATATCGACAGCAACAGAAATCACTGCGCAAT 2494
QY 2701 CAATACTTGGTACCAATTTCAACCTCTGGCAACTGCAAGCTTGGTGTGTAACAGCAT 2760
Db 2495 CAATACTTGGTACCAATTTCAACCTCTGGCAACTGCAAGCTTGGTGTGTAACAGCAT 2554
QY 2761 TATCAGCAATTAACATAGAACCTTCAAGGCGCCAGAGCATGTCAATTAATCAACAGGCTG 2820
Db 2555 TATCAGCAATTAACATAGAACCTTCAAGGCGCCAGAGCATGTCAATTAATCAACAGGCTG 2614
QY 2821 ATGTGTCAATTTGAACCATCATCTTGGCAGCAGTAACGAACTGTTTCTCGGCAAAATA 2880
Db 2615 ATGTGTCAATTTGAACCATCATCTTGGCAGCAGTAACGAACTGTTTCTCGGCAAAATA 2674
QY 2881 TTTTGTCACTTTAGGAACAGTTTCTGTTCCGTAAGCCATAGCAGTCTACCTCGGCGGCG 2940
Db 2675 TTTTGTCACTTTAGGAACAGTTTCTGTTCCGTAAGCCATAGCAGTCTACCTCGGCGGCG 2734
QY 2941 CTCCTGCTAGCAGCATACATTTAGCAACCACTTGTGGCAACGATAGATCACTTCTGGGG 3000
Db 2735 CTCCTGCTAGCAGCATACATTTAGCAACCACTTGTGGCAACGATAGATCACTTCTGGGG 2794
QY 3001 TAAGGTTACCATCTCTTCTAGTGGAGATGCAAAAAAATTTCTTTGCAACAGCAACTT 3060
Db 2795 TAAGGTTACCATCTCTTCTAGTGGAGATGCAAAAAAATTTCTTTGCAACAGCAACTT 2854
QY 3061 TGGCAGGAACACCCAGCATCAGGGAAGTGGGAAGCAGAAATTTGCGGTTCCACAGGAATAT 3120
Db 2855 TGGCAGGAACACCCAGCATCAGGGAAGTGGGAAGCAGAAATTTGCGGTTCCACAGGAATAT 2914

QY 3121 AGAGGCCAACTTTCTCAATAGTCTTGTCAAAAACGAGAGCAGACTACACAGGGCAAGTCT 3180
Db 2915 AGAGGCCAACTTTCTCAATAGTCTTGTCAAAAACGAGAGCAGACTACACAGGGCAAGTCT 2974
QY 3181 CAACCTTGCAACGCTCTCGTTAGTTAGCTTCATGGAAATTTCTCGACGTTATCTATAGAGA 3240
Db 2975 CAACCTTGCAACGCTCTCGTTAGTTAGCTTCATGGAAATTTCTCGACGTTATCTATAGAGA 3034
QY 3241 GATCAATGGCTCTCTTAACGTTATCTGGCAATTCATAAAGTTCTCTGGGAAAGGAGCTT 3300
Db 3035 GATCAATGGCTCTCTTAACGTTATCTGGCAATTCATAAAGTTCTCTGGGAAAGGAGCTT 3094
QY 3301 CTAAACACAGGTTCTCTTAACGCGACTCCATCAAACTTTGGCAGTTAGTTCTTAAAGGGCTT 3360
Db 3095 CTAAACACAGGTTCTCTTAACGCGACTCCATCAAACTTTGGCAGTTAGTTCTTAAAGGGCTT 3154
QY 3361 TGTCAACCAATTTTGACGACATTTGTCGAATTTGTTGACTAATTCATAATCTGTTCCG 3420
Db 3155 TGTCAACCAATTTTGACGACATTTGTCGAATTTGTTGACTAATTCATAATCTGTTCCG 3214
QY 3421 TTTTCTGATAGGACGAGGAGCATCTTCAATTTCTGTGAGGAGGCTTTAGAAAGCT 3480
Db 3215 TTTTCTGATAGGACGAGGAGCATCTTCAATTTCTGTGAGGAGGCTTTAGAAAGCT 3274
QY 3481 CAATTTTGCACAATTCATAGACATTTGTCGAAGGAGCTTCTTTAGGTTTGGATTTCTTCT 3540
Db 3275 CAATTTTGCACAATTCATAGACATTTGTCGAAGGAGCTTCTTTAGGTTTGGATTTCTTCT 3334
QY 3541 TAGGTTGTTCTTGGTGATCTTGGCTTGGCATCTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3600
Db 3335 TAGGTTGTTCTTGGTGATCTTGGCTTGGCATCTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3394
QY 3601 ACTTCATATCAGGTTTCTTCTCCTGCTTGGCATCTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3660
Db 3395 ACTTCATATCAGGTTTCTTCTCCTGCTTGGCATCTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3454
QY 3661 CTAAATGCAAAATAAATAAGTCAACATTTCCAGGCTTATCTTCTTCTTCTTCTTCTTCTTCTT 3720
Db 3455 CTAAATGCAAAATAAATAAGTCAACATTTCCAGGCTTATCTTCTTCTTCTTCTTCTTCTTCTT 3514
QY 3721 CTGCAAGTTCACTAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3780
Db 3515 CTGCAAGTTCACTAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3574
QY 3781 ACCGTTTGGTATAGAACTTCTGAGCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3840
Db 3575 ACCGTTTGGTATAGAACTTCTGAGCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3634
QY 3841 TGGCTCTAAGACCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3900
Db 3635 TGGCTCTAAGACCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3694
QY 3901 CACCTGTTTGTTCACCAAAATTTCAAGCAGTCTCCATCAATCCAAATCCAAATCCAAATCCAAATCCAAAT 3960
Db 3695 CACCTGTTTGTTCACCAAAATTTCAAGCAGTCTCCATCAATCCAAATCCAAATCCAAATCCAAATCCAAAT 3794
QY 3961 GCAACTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 4020
Db 3755 GCAACTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3814
QY 4021 GGTAGACTCCAGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 4080
Db 3815 GGTAGACTCCAGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3874
QY 4081 GSCCAACGAGTAATTAAGAGAGTCAGCCACCAAAAGTGTGAATAGACCAATTCGGGGGGGT 4140
Db 3875 GSCCAACGAGTAATTAAGAGAGTCAGCCACCAAAAGTGTGAATAGACCAATTCGGGGGGGT 3994
QY 4141 CAGTGTCAAGACGCCCAAAATTTCACTGACAGGGAACTTTTGTGACATCTTTCAGNAA 4200
Db 3935 CAGTGTCAAGACGCCCAAAATTTCACTGACAGGGAACTTTTGTGACATCTTTCAGNAA 3994
QY 4201 GTTCTGTTTCTAGTAGTCAATTTGCCGAGCATCAATAATGGGGATTATACCAGAGCAACAG 4260

Db 3995 GTTCGTATTCAGTAGTCATATTCGCGAGCATCAATATGCGGATTTATACCAAGCAACAG 4054
Qy 4261 TGGAAAGTCACATCTACCAACTTTTCGGTCTCAGAAAAAGCATAAACAGTTCTACTACCGC 4320
Db 4055 TGGAAAGTCACATCTACCAACTTTTCGGTCTCAGAAAAAGCATAAACAGTTCTACTACCGC 4114
Qy 4321 CATAGTGAAACTTTTCAAAATCGCCAGTGGAGAAAGAAAGGCACACGCGATCTAGCAT 4380
Db 4115 CATTAGTGAAGTCTTTCAATTCGCCAGTGGAGAAAGAAAGGCACACGCGATCTAGCAT 4174
Qy 4381 TAGCGGCAAGATGCAACTTTTATCAACCGAGTCTTATAGATAACCCCTAGCGCTGGGA 4440
Db 4175 TAGCGGCAAGATGCAACTTTTATCAACCGAGTCTTATAGATAACCCCTAGCGCTGGGA 4234
Qy 4441 TCATCCCTTGGACAACCTTTCTGCCAAATCTAGTCCAAATCACTTCATTTGATACCAT 4500
Db 4235 TCATCCCTTGGACAACCTTTCTGCCAAATCTAGTCCAAATCACTTCATTTGATACCAT 4294
Qy 4501 TATTGTACAACTTGAGCAAGTGTTCGATCAGTCTCTCAAAATGGTCTCTGTAAACGGATG 4560
Db 4295 TATTGTACAACTTGAGCAAGTGTTCGATCAGTCTCTCAAAATGGTCTCTGTAAACGGATG 4354
Qy 4561 ACTCAACTTGCATTAACCTTGAAGCTCAGTCGATTCAGTGAACCTTGATCAGGTTGTGCA 4620
Db 4355 ACTCAACTTGCATTAACCTTGAAGCTCAGTCGATTCAGTGAACCTTGATCAGGTTGTGCA 4414
Qy 4621 GCTGTCAGCAGCATAGGAAACACCGCTTTCTACCAACTCAAGAAATTAACAACCT 4680
Db 4415 GCTGTCAGCAGCATAGGAAACACCGCTTTCTACCAACTCAAGAAATTAACAACCT 4474
Qy 4681 CTGCAACACTTTCGCTATGACAGTACAGGAAATGTATCACTTTGAAGTGGACAGTGAG 4740
Db 4475 CTGCAACACTTTCGCTATGACAGTACAGGAAATGTATCACTTTGAAGTGGACAGTGAG 4534
Qy 4741 TGTAGTCTTGAGAAATCTGAGCGGTATTTTATATCATCAGTGAAGTCAATCAGGAGA 4800
Db 4535 TGTAGTCTTGAGAAATCTGAGCGGTATTTTATATCATCAGTGAAGTCAATCAGGAGA 4594
Qy 4801 TCCTCTAGCGGAGCATCGTGGCGCA 4828
Db 4595 TCCTCTAGCGGAGCATCGTGGCGCACTGACGGGGGGGGGGGGGGGGTCTGC 4654
Qy 4829 4828
Db 4655 CTCGTGAAGAAGTGTTCGTGACTCATACAGGCTGAATCGCCCATCATCCAGCCAGA 4714
Qy 4829 4828
Db 4715 AAGTGAGGAGCCACGGTTGATGAGAGCTTTGTTGTTAGTGACAGTGGTGATTTGA 4774
Qy 4829 4828
Db 4775 ACTTTTGTCTTCCACGGAACGGTCTGCTGTTGTCGGGAAGATGCGTGATCTGATCCTTCA 4834
Qy 4829 4828
Db 4835 ACTCAGAAAAGTTCGATTTTATCAAAAGCGCGCGTCCCGTCAAGTCAGCGTAATGCT 4894
Qy 4829 4828
Db 4895 CTGCCAGTGTACAAACCAATTAACCAATTCGTATTAAGAAAACATCATCGAGCATCAAAATG 4954
Qy 4829 4828
Db 4955 AAACCTGCAATTTTATCATATCAGGATTAATCAATACCATATTTTGAAGAAACCGTTCTG 5014
Qy 4829 4828
Db 5015 TAATGAAGGAGAAAACCTACCGAGCGAGTTCATAGGATGCAAGATCCTCGTATCGGTC 5074
Qy 4829 4828

Db 5075 TGGATTCCGACTCGTCCAAACATCAATACAACTTATTAATTTCCCTCCTCAAAATAAG 5134
Qy 4829 4828
Db 5135 GTTATCAAGTGAGAAAATCAACATGAGTACGACTGAATCCGGTGAGAAATGGCAAAAGCTT 5194
Qy 4829 4828
Db 5195 ATGCATTTCTTTCCAGACTTGTTCACAGGCCAGCCATTACGCTCGTCATCAAAATCACT 5254
Qy 4829 4828
Db 5255 CGCATCAACCAAAACCGTTTATCAATTCGTGATTCGCGCTGAGCGAGACGAAATACCGCATC 5314
Qy 4829 4828
Db 5315 GCTGTTAAAGAGCAATTAACAAACAGGATCGAATGCAACCGCGCAGGAAACACTGCCAG 5374
Qy 4829 4828
Db 5375 CGCATCAACATATTTTACCTGAATCAGGATATTTCTTAATACCTGGAATGCTGTTTT 5434
Qy 4829 4828
Db 5435 CCGGGGATCGCAGTGGTGAGTAACCATGTCATCATCAGGAGTACGATTAATAATGTTGAT 5494
Qy 4829 4828
Db 5495 GGTGGAGAGGAGATAAATTCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATC 5554
Qy 4829 4828
Db 5555 ATTGGCAACGCTACCTTTGCCATGTTTCAGAAAACAACTCTGGCGCATCGGCTTCCCAT 5614
Qy 4829 4828
Db 5615 CAATCGATAGATTGTCGCACCTGATTGCCGACATTCGCCGACATTCGCCGCTTATACCAT 5674
Qy 4829 4828
Db 5675 TAAATCAGCATCCATGTTGGAATTTAATCGCGCCTCGAGCAAGAGTTTCCGCTTGAAT 5734
Qy 4829 4828
Db 5735 ATGGCTCATACACCCCTTGTATTTACTGTTTATGTAAGCAGACAGTTTATTTGTTTCATGA 5794
Qy 4829 4828
Db 5795 TGATAATATTTTATCTTGTGCAATGTACATCAGAGATTTTGAGACACAACTGGCTTTC 5854
Qy 4829 4828
Db 5855 CCCCCCCCCCTGTCAGGTGCGCATCACCGCGCCACAGTGGCTTGTGGCGCTTATAT 5914
Qy 4880 CGCCGACATCACCGATGGGAAAGATCGGCTCGCCACTTTCGGGCTCATGAGCGCTTGT 4939
Db 5915 CGCCGACATCACCGATGGGAAAGATCGGCTCGCCACTTTCGGGCTCATGAGCGCTTGT 5974
Qy 4940 CGCGTGGGTATGTTGGCAGGCCCCCGTGGCCGGGAGTGTGGCGGCTATCTCTTGA 4999
Db 5975 CGCGTGGGTATGTTGGCAGGCCCCCGTGGCCGGGAGTGTGGCGGCTATCTCTTGA 6034
Qy 5000 TGCACCATTTCTTGGCGCGCGTGTCTCAACCGCCTCAACCTACTCTGGGCTGCTTCT 5059
Db 6035 TGCACCATTTCTTGGCGCGCGTGTCTCAACCGCCTCAACCTACTCTGGGCTGCTTCT 6094
Qy 5060 AATGCAAGGAGTCCGATTAAGGAGAGCGTTCGAGTATCTATGATGGAAGTATGGGAATGTT 5119
Db 6095 AATGCAAGGAGTCCGATTAAGGAGAGCGTTCGAGTATCTATGATGGAAGTATGGGAATGTT 6154
Qy 5120 GATACCGCATTTCTTCAAGTGTCTTGGGCTCTCTATCAGATTCCTCAACCTAAAGCAAC 5179
Db 6155 GATACCGCATTTCTTCAAGTGTCTTGGGCTCTCTATCAGATTCCTCAACCTAAAGCAAC 6214

QY	5180	CGAGGAGGAGATTTCAATGTTAAATTTCTCTGACTTTTGGTCAATCAAGTACATCGAACTG	5239	QY	6260	CAAAAGCCAGCAAAAGCCAGGAAACCGTAAAGAGCCGCGTTCCTCGCGTTTTCCTCAT	6319
DB	6215	CGAGGAGGAGATTTCAATGTTAAATTTCTCTGACTTTTGGTCAATCAAGTACATCGAACTG	6274	DB	7295	CAAAAGCCAGCAAAAGCCAGGAAACCGTAAAGAGCCGCGTTCCTCGCGTTTTCCTCAT	7354
QY	5240	TGAGACTATCTCGGTATGACAGCAAAATGTCCTCTTCGGAGACAGTAAATGAAGTCCC	5299	QY	6320	GGCTCGCGCCCTCTGACGAGCATCAAAAAATCGACGCTCAAGTCAAGGTGCGGAAACC	6379
DB	6275	TGAGACTATCTCGGTATGACAGCAAAATGTCCTCTTCGGAGACAGTAAATGAAGTCCC	6334	DB	7355	GGCTCGCGCCCTCTGACGAGCATCAAAAAATCGACGCTCAAGTCAAGGTGCGGAAACC	7414
QY	5300	ACCAATAAGAAATCTTGTGTTATACGGAACAAATCTTCTTTTCGAACTTTTTCGGTGGC	5359	QY	6380	CGACAGGACTATAAGATACCAAGCGCTTTCCCTCGGAAGTCCCTCGTGGCTCTCTCTG	6439
DB	6335	ACCAATAAGAAATCTTGTGTTATACGGAACAAATCTTCTTTTCGAACTTTTTCGGTGGC	6394	DB	7415	CGACAGGACTATAAGATACCAAGCGCTTTCCCTCGGAAGTCCCTCGTGGCTCTCTCTG	7474
QY	5360	TTGAACATATAAGTATGATGATATGTCGGGTAGGAAATGGAAGCGGCAAAATGCTTACC	5419	QY	6440	TTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGC	6499
DB	6395	TTGAACATATAAGTATGATGATATGTCGGGTAGGAAATGGAAGCGGCAAAATGCTTACC	5454	DB	7475	TTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGC	7534
QY	5420	TTCTGGACCTTCAAGAGGTATGTAGGTTTGTAGATCTGATGCCAACTTCAGTGAACAAC	5479	QY	6500	TTTCTCATAGTCTACGCTGTAGGTATCTCAGTTTCGGTGTAGTTCGTTCCGCTCAAGCTGG	6559
DB	6455	TTCTGGACCTTCAAGAGGTATGTAGGTTTGTAGATCTGATGCCAACTTCAGTGAACAAC	6514	DB	7535	TTTCTCATAGTCTACGCTGTAGGTATCTCAGTTTCGGTGTAGTTCGTTCCGCTCAAGCTGG	7594
QY	5480	GTTGCTATTTGCTTCAACCAATCCGAATCCAGAGAAATCAAAAGTTGTTTGTCTACTAT	5539	QY	6560	GCTGTGTGCAAGAAACCCCGCTTCAGCCCGACCGCTTCGCTTATCCGGTAACTATCGTC	6619
DB	6515	GTTGCTATTTGCTTCAACCAATCCGAATCCAGAGAAATCAAAAGTTGTTTGTCTACTAT	6574	DB	7595	GCTGTGTGCAAGAAACCCCGCTTCAGCCCGACCGCTTCGCTTATCCGGTAACTATCGTC	7654
QY	5540	GATCCAGCCAGTCCGCTTGTGAACCTGCAATAGTGTCTGCTGTTTTCGAGTCTATCT	5599	QY	6620	TTGATCCAAACCCCGTAAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGA	6679
DB	6575	GATCCAGCCAGTCCGCTTGTGAACCTGCAATAGTGTCTGCTGTTTTCGAGTCTATCT	6634	DB	7655	TTGATCCAAACCCCGTAAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGA	7714
QY	5600	TGATGAATAAATCTAGTCTTTGTATTAATAATCTTGAAGCAAGCGGAGTAATATACC	5659	QY	6680	TTAGCAGAGCAGGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGGTGGCTTAACATACG	6739
DB	6635	TGATGAATAAATCTAGTCTTTGTATTAATAATCTTGAAGCAAGCGGAGTAATATACC	6694	DB	7715	TTAGCAGAGCAGGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGGTGGCTTAACATACG	7774
QY	5660	CABATCTAAACCTCTTTTAAACGTTAAAGGACAGTATGTCCTGCTGTATTAAACCCC	5719	QY	6740	GCTACACTAGAAAGACAGTATTTGGTATCTGCTGCTGCTGAAGACAGTACCTTCGGAA	6799
DB	6695	CAATCTAAACCTCTTTTAAACGTTAAAGGACAGTATGTCCTGCTGTATTAAACCCC	6754	DB	7775	GCTACACTAGAAAGACAGTATTTGGTATCTGCTGCTGCTGAAGACAGTACCTTCGGAA	7834
QY	5720	AAATCAGCTCTAGTCTGATCCTCATCAACTTTGAGGGGCACTATCTTGTTTTATAGAAAT	5779	QY	6800	AAAGAGTTGGTAGCTCTTGATCCGCAACAAACCAACCGCTGGTAGCGGTGTTTTTTTG	6859
DB	6755	AAATCAGCTCTAGTCTGATCCTCATCAACTTTGAGGGGCACTATCTTGTTTTATAGAAAT	6814	DB	7835	AAAGAGTTGGTAGCTCTTGATCCGCAACAAACCAACCGCTGGTAGCGGTGTTTTTTTG	7894
QY	5780	TTGCGAGATCGATATCGAGAAAGGTAGCTGATTTTAAACGTTGATTTTATCTCAA	5839	QY	6860	TTTTCGACAGCAGATTAACGCGCAAGAAAGAGTCTCAAGAAAGTCCCTTTGATCTTTT	6919
DB	6815	TTGCGAGATCGATATCGAGAAAGGTAGCTGATTTTAAACGTTGATTTTATCTCAA	6874	DB	7895	TTTTCGACAGCAGATTAACGCGCAAGAAAGAGTCTCAAGAAAGTCCCTTTGATCTTTT	7954
QY	5840	GATCTCTGCTCGCGCTTTTCGGTGTATGACGGTGAACCTCTGACATGACGCTCCCG	5899	QY	6920	CTACGCGGTCTGACGCTCAGTGGAAACGAAACTCACCTTAAAGGATTTTGGTCAATGAGAT	6979
DB	6875	GATCTCTGCTCGCGCTTTTCGGTGTATGACGGTGAACCTCTGACATGACGCTCCCG	6934	DB	7955	CTACGCGGTCTGACGCTCAGTGGAAACGAAACTCACCTTAAAGGATTTTGGTCAATGAGAT	8014
QY	5900	GAGACGCTCAGCTGTTCTGTAAAGCGATGTCGGGAGGAGACAGCCGCTCAGGCGCG	5959	QY	6980	TATCAAAAGGATCTTCACTTAGATCTTTTAAATTAAGTCTTTTAAATCAATCT	7039
DB	6935	GAGACGCTCAGCTGTTCTGTAAAGCGATGTCGGGAGGAGACAGCCGCTCAGGCGCG	6994	DB	8015	TATCAAAAGGATCTTCACTTAGATCTTTTAAATTAAGTCTTTTAAATCAATCT	8074
QY	5960	TCAGCGGTGTTGCGGCTGTCGGGCGCAGCCATGACCCAGTCACTAGCGATAGCGGA	6019	QY	7040	AAAGTATATATGATTAACCTTGGTCTGACATTAACGCTTAATCAGTGAAGCACTTA	7099
DB	6995	TCAGCGGTGTTGCGGCTGTCGGGCGCAGCCATGACCCAGTCACTAGCGATAGCGGA	7054	DB	8075	AAAGTATATATGATTAACCTTGGTCTGACATTAACGCTTAATCAGTGAAGCACTTA	8134
QY	6020	GTGATATGCTTAACTATGCGGCATCAGAGCAGATTTGATCTGAGTGCACATATGTC	6079	QY	7100	TCTCAGCATCTGTCTATTTCTGTTTCTATCTCATCTAGTTCGCTGCTCCCGTCTGATATAA	7159
DB	7055	GTGATATGCTTAACTATGCGGCATCAGAGCAGATTTGATCTGAGTGCACATATGTC	7114	DB	8135	TCTCAGCATCTGTCTATTTCTGTTTCTATCTCATCTAGTTCGCTGCTCCCGTCTGATATAA	8194
QY	6080	GGTGTGAATACCCACAGATGCTTAAAGGAGAAATACCGCATCAGCGCTTCCGCTT	6139	QY	7160	CTACGATACGGAGGCTTACCATCTGCGCCAGTGTGCTGCAATGATACCGGAGACCCAC	7219
DB	7115	GGTGTGAATACCCACAGATGCTTAAAGGAGAAATACCGCATCAGCGCTTCCGCTT	7174	DB	8195	CTACGATACGGAGGCTTACCATCTGCGCCAGTGTGCTGCAATGATACCGGAGACCCAC	8254
QY	6140	CCTCGCTCACTGACTCGCTCGCTCGCTCGTTCGGCTCGCGGAGCGGTATACGCTCACT	6199	QY	7220	GCTCACCGCTCCAGATTTATCAGCAATAAACCCAGCCAGCGGAGGCGGAGGCGAGAA	7279
DB	7175	CCTCGCTCACTGACTCGCTCGCTCGCTCGTTCGGCTCGCGGAGCGGTATACGCTCACT	7234	DB	8255	GCTCACCGCTCCAGATTTATCAGCAATAAACCCAGCCAGCGGAGGCGGAGGCGAGAA	8314
QY	6200	CAAAAGCGGTAAATAGGTTATCCACAGATCAGGGGATTAACGAGGAAACATGTTAG	6259	QY	7280	GTGTCCTCGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAACTAGAG	7339
DB	7235	CAAAAGCGGTAAATAGGTTATCCACAGATCAGGGGATTAACGAGGAAACATGTTAG	7294	DB	8315	GTGTCCTCGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAACTAGAG	8374
				QY	7340	TAAGTAGTTCGCGAGTTAAATAGTTTGGCAACGTTGTTGCCATTGCTGCGGATCTGTTG	7399


```

8375 TAAGTAGTTCGCGAGTTAATAGTTTCGCGAAGGTTGTTGCCATTGCTCGAGCATCGTGG 8434
8400 TGTCAAGTTCGCGTTCGTTGGTATGCTTCAATCAGCTCGGTTCCCAAGATCAAGGCGAG 7459
8435 TGTCAAGTTCGCGTTCGTTGGTATGCTTCAATCAGCTCGGTTCCCAAGATCAAGGCGAG 8494
8460 TTACATGATCCCAATGTTGCGAAGGTTAGCTCCTCGGTCCTCGATCGTTG 7519
8495 TTACATGATCCCAATGTTGCGAAGGTTAGCTCCTCGGTCCTCGATCGTTG 8554
8520 TCAGAGTAAGTTGGCGAGTGTATCACTCATGTTATGGCAGCATGATTAATCTC 7579
8555 TCAGAGTAAGTTGGCGAGTGTATCACTCATGTTATGGCAGCATGATTAATCTC 8614
8580 TTACTGTCTATGCGATCGTAAGATGCTTTTCTGTGACTGGTGAGTCAACCAAGTCAT 7639
8615 TTACTGTCTATGCGATCGTAAGATGCTTTTCTGTGACTGGTGAGTCAACCAAGTCAT 8674
8640 TCTGAGATAGTGTGCGGCGACCGAGTGTCTTTCGCGCGTCAACCGGATATA 7699
8675 TCTGAGATAGTGTGCGGCGACCGAGTGTCTTTCGCGCGTCAACCGGATATA 8734
8700 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCAATGGAAAGCTTTCTTCGGGGCGAA 7759
8735 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCAATGGAAAGCTTTCTTCGGGGCGAA 8794
8760 AACTCTCAAGATCTTACCGCTGTGAGATCCAGTTCGATTAACCCACTCGTCACCCA 7819
8795 AACTCTCAAGATCTTACCGCTGTGAGATCCAGTTCGATTAACCCACTCGTCACCCA 8854
8820 ACTGATCTTCAGCATCTTTTACTTTTCAACAGCGTTTCTGGGTGACCAAAACAGGAAGC 7879
8855 ACTGATCTTCAGCATCTTTTACTTTTCAACAGCGTTTCTGGGTGACCAAAACAGGAAGC 8914
8880 AAAATGCCCAAAAAGGGAATAGGCGGACACCGGAATGTTGAATATCTCATCTCTCC 7939
8915 AAAATGCCCAAAAAGGGAATAGGCGGACACCGGAATGTTGAATATCTCATCTCTCC 8974
8940 TTTTTCATATTTAAGCACTTTATCAGGTTTATGCTCATGAGCGGATACATATTG 7999
8975 TTTTTCATATTTAAGCACTTTATCAGGTTTATGCTCATGAGCGGATACATATTG 9034
8000 AATGTATTTAGAAAATAAACAATAGGGGTTCCGCGCATATTTCCCGAAAAGTGCCAC 8059
9035 AATGTATTTAGAAAATAAACAATAGGGGTTCCGCGCATATTTCCCGAAAAGTGCCAC 9094
8060 CTGAGCTCTAAGAACCATTTATATCATGACATTAACCTTATAAATAGGCGTATCAGA 8119
9095 CTGAGCTCTAAGAACCATTTATATCATGACATTAACCTTATAAATAGGCGTATCAGA 9154
8120 GGCCTTTCTGCTTTCAGAAATTAATCTCATGTTTGACAGCTTATCATCGATAAGCTGAC 8179
9155 GGCCTTTCTGCTTTCAGAAATTAATCTCATGTTTGACAGCTTATCATCGATAAGCTGAC 9214
8180 TATGTTGTTGTTGTAATAGACGAGATCGGGAACACTGAAATTAACAGTTATTATT 8239
9215 TCATGTTGTTGTTGTAATAGACGAGATCGGGAACACTGAAATTAACAGTTATTATT 9274
8240 CG 8241
9275 CG 9276

RESULT 4
US-10-267-219-13
; Sequence 13, Application US/10267219
; Publication No. US20030143219A1
; GENERAL INFORMATION:
; APPLICANT: Madison, Edwin
; APPLICANT: Yeh, Jium-Chern
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING A TRANSMEMBRANE SERINE PROTEASE 2
; TITLE OF INVENTION: ENCODED POLYPEPTIDES AND METHODS BASED THEREON
```

```

; FILE REFERENCE: 24745-1621
; CURRENT APPLICATION NUMBER: US/10/267,219
; CURRENT FILING DATE: 2002-10-08
; PRIOR APPLICATION NUMBER: 60/328,530
; PRIOR FILING DATE: 09-OCT-2001
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 9276
; TYPE: DNA
; ORGANISM: Pichia pastoris
; US-10-267-219-13
```

```

Query Match      79.6%; Score 6561.6; DB 15; Length 9276;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 803; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;

Qy      1 AGATCTAACATCCAAAGAGGTTGAATGAACCTTTTCCCATCCGACATCCACAG 60
Db      1 AGATCTAACATCCAAAGAGGTTGAATGAACCTTTTCCCATCCGACATCCACAG 60

Qy      61 GTCCATCTCACATAGTCCAAAGCGCAACAGGAGGGGATACACTAGCAGCAGACCGT 120
Db      61 GTCCATCTCACATAGTCCAAAGCGCAACAGGAGGGGATACACTAGCAGCAGACCGT 120

Qy      121 TGCAAGCGCAGGACCTCCACTCTCTCTCTCAACACCCACCTTTTGCCATCGAAAACC 180
Db      121 TGCAAGCGCAGGACCTCCACTCTCTCTCTCAACACCCACCTTTTGCCATCGAAAACC 180

Qy      181 AGCCAGTTATGGGCTTGATGGAGTCTGCTCATTCCTCAATCTCTCTATAGGCTACTA 240
Db      181 AGCCAGTTATGGGCTTGATGGAGTCTGCTCATTCCTCAATCTCTCTATAGGCTACTA 240

Qy      241 ACACCATGCTTTATTAGCCTGCTATCTGCGGCCCTCTGCGAGGTTTCATGTTGTTTA 300
Db      241 ACACCATGCTTTATTAGCCTGCTATCTGCGGCCCTCTGCGAGGTTTCATGTTGTTTA 300

Qy      301 TTTCCGAATCAACAAGCTCCGATTAACCCGAAATCACTCCAGATGAGGCTTTCTG 360
Db      301 TTTCCGAATCAACAAGCTCCGATTAACCCGAAATCACTCCAGATGAGGCTTTCTG 360

Qy      361 AGTGTGGGTCAATAGTTTCATGTTCCCAATGCGCCCAAACTGACAGTTTAAACGCT 420
Db      361 AGTGTGGGTCAATAGTTTCATGTTCCCAATGCGCCCAAACTGACAGTTTAAACGCT 420

Qy      421 GTCTTGGAACTAATATGACAAAGCGTGATCTCATCCAAAGATGAATAGTTTGGTTCG 480
Db      421 GTCTTGGAACTAATATGACAAAGCGTGATCTCATCCAAAGATGAATAGTTTGGTTCG 480

Qy      481 TTGAATGCTAAACCGCGAGTTGGTCAAAAGAACTTCCAAAGTCCCATACCGTTGT 540
Db      481 TTGAATGCTAAACCGCGAGTTGGTCAAAAGAACTTCCAAAGTCCCATACCGTTGT 540

Qy      541 CTTGTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAATGCTTAGGCGAGTCT 600
Db      541 CTTGTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAATGCTTAGGCGAGTCT 600

Qy      601 CTCTATCGCTTCTGAACCCCGGTGACCTGTGCGGAAACGCAAAATGGGGGAAACACCCGCT 660
Db      601 CTCTATCGCTTCTGAACCCCGGTGACCTGTGCGGAAACGCAAAATGGGGGAAACACCCGCT 660

Qy      661 TTTTGGATGATTATGCAATGCTCTCCCATGTTGCTTCCAAAGATTCCTGCGGAATACT 720
Db      661 TTTTGGATGATTATGCAATGCTCTCCCATGTTGCTTCCAAAGATTCCTGCGGAATACT 720

Qy      721 GCTGATAGCCTTAACGTTTCATGATCAAAAATTAATGTTCTAACCCCTTACTTGACAGCAAT 780
Db      721 GCTGATAGCCTTAACGTTTCATGATCAAAAATTAATGTTCTAACCCCTTACTTGACAGCAAT 780

Qy      781 ATATAACAGAGGAGGAGCTGCCCTGCTTAAACCTTTTTTTTATCATCATTTATTAGCTT 840
Db      781 ATATAACAGAGGAGGAGCTGCCCTGCTTAAACCTTTTTTTTATCATCATTTATTAGCTT 840
```

841 ACTTTCATAATGGACGTGGTTCCAAATGACAAGCTTTGATTTTAAACGACCTTTTAAACGA 900
841 ACTTTCATAATGGACGTGGTTCCAAATGACAAGCTTTGATTTTAAACGACCTTTTAAACGA 900
901 CAACCTTGAGAGATCAAAAAACAATAATATTCGAAGATCCAAACGATGAGATTTCCCT 960
901 CAACCTTGAGAGATCAAAAAACAATAATATTCGAAGATCCAAACGATGAGATTTCCCT 960
961 TCAATTTTACTGCAAGTTTATTCGCAGCATCTCCGCAATAGCTGCTCCAGTCAACACT 1020
961 TCAATTTTACTGCAAGTTTATTCGCAGCATCTCCGCAATAGCTGCTCCAGTCAACACT 1020
1021 ACAACAGAAGATGAACCGGCAAAATTCGGCTGAAGCTGTCATCGGTTACTCAGATTTA 1080
1021 ACAACAGAAGATGAACCGGCAAAATTCGGCTGAAGCTGTCATCGGTTACTCAGATTTA 1080
1081 GAAGGGGATTTGATGTTGCTGTTTGGCATTTTCCACAGCACAATAACGGGTTATTG 1140
1081 GAAGGGGATTTGATGTTGCTGTTTGGCATTTTCCACAGCACAATAACGGGTTATTG 1140
1141 TTTATAAATACTACTATTTGCCAGCATTCGCTGCTAAAGAAAGGGGATCTCTCGAGAAA 1200
1141 TTTATAAATACTACTATTTGCCAGCATTCGCTGCTAAAGAAAGGGGATCTCTCGAGAAA 1200
1201 AGAGAGGCTGAAGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCC 1260
1201 AGAGAGGCTGAAGCTTA----- 1217
1261 GCCCCCCGGAGAGGATAAATGTGGCTTCCCGGTGTCACCGCCAGCAGTGCACGGAG 1320
1218 ----- 1217
1321 AGAGGTTGCTGTTTGTGATGACAGTGTCCGGGATTCGCGTGGTCTCCACCCCATGGCC 1380
1218 ----- 1217
1381 ATGGAGACACTCAAGAGAGAAGATGTCCTTCTAACTAGTGGCGTAGAATTCCTTAGGG 1440
1218 -----CGTAGAATTCCTTAGGG 1234
1441 CGGCCGGAATTAATTCGCCTTAGACATGACTGTTCTCAGTTCAGTTCAGGCTTAGG 1500
1235 CGGCCGGAATTAATTCGCCTTAGACATGACTGTTCTCAGTTCAGTTCAGGCTTAGG 1294
1501 AGAGACCGGCTTCTGCTAGATCTAAATCAAGAGATGTCAGATGCCATTTGCTGAGAG 1560
1295 AGAGACCGGCTTCTGCTAGATCTAAATCAAGAGATGTCAGATGCCATTTGCTGAGAG 1354
1561 ATGCAGGCTTCATTTTTCATATCTTTTATTTTGTAACTATATAGTATAGGATTTTTCCT 1620
1355 ATGCAGGCTTCATTTTTCATATCTTTTATTTTGTAACTATATAGTATAGGATTTTTCCT 1414
1621 GTCAATTTGTTCTTCTCGTAGAGCTGCTGCTGATCAGGCTATCTCGCAGCTGATGAA 1680
1415 GTCAATTTGTTCTTCTCGTAGAGCTGCTGCTGATCAGGCTATCTCGCAGCTGATGAA 1474
1681 TATCTTTGTTGAGGGTTTGGGAAAATCATTCGAGTTTGTATGTTTCTTTGTTATTTCCC 1740
1475 TATCTTTGTTGAGGGTTTGGGAAAATCATTCGAGTTTGTATGTTTCTTTGTTATTTCCC 1534
1741 ACTCCTCTTCAGAGPACAGATTAAGTGAAGTGTGGTTTGTGCAAGCTTATPCGATAA 1800
1535 ACTCCTCTTCAGAGPACAGATTAAGTGAAGTGTGGTTTGTGCAAGCTTATPCGATAA 1594
1801 GCTTTAAATTCGGTAGTTTATCACAGTTTAAATTCGTAAAGCAGTCCAGCCGCTGATGAA 1860
1595 GCTTTAAATTCGGTAGTTTATCACAGTTTAAATTCGTAAAGCAGTCCAGCCGCTGATGAA 1654
1861 ATCTAACCAATGGCTTCATGTCATCTCGGACCGCTCAACCTGGATGCTGTAGGCATPAGG 1920
1655 ATCTAACCAATGGCTTCATGTCATCTCGGACCGCTCAACCTGGATGCTGTAGGCATPAGG 1714
1921 CTGTTGTTATGCGGTACTGCGCGGGCTCTTGGGGATATCGTCCATTCGCCAGCATCGC 1980

1715 CTTGGTTATGCCGGTACTGCGGGCTCTTGGGATATCGTCCATTCGACAGCATCGC 1774
1981 CAGTCACATATGCGGTGCTGCTAGCGCTATATGCGTTGATGCAATTTCTATCGCACCCGT 2040
1775 CAGTCACATATGCGGTGCTGCTAGCGCTATATGCGTTGATGCAATTTCTATCGCACCCGT 1834
2041 TCTCGAGACATGCTCGGACCGCTTTGGCGCGCCGACGCTCTGCTCGCTCTGCTACTTGG 2100
1835 TCTCGAGACATGCTCGGACCGCTTTGGCGCGCCGACGCTCTGCTCGCTCTGCTACTTGG 1894
2101 AGCCACTATCGACTAGCGCATCGGACGACACACCCGCTCTGCTGATCTCAATCTA 2160
1895 AGCCACTATCGACTAGCGCATCGGACGACACACCCGCTCTGCTGATCTCAATCTA 1954
2161 AATGTAAGTTAAATCTCTAAATAATTAATTAAGTCCAGGTTTCTCCATACGAACCTTAA 2220
1955 AATGTAAGTTAAATCTCTAAATAATTAATTAAGTCCAGGTTTCTCCATACGAACCTTAA 2014
2221 CAGCATTCGCGTGAAGATCTAGACCTTCAACAGAGCCAGATCCATCTACTGCTGGCCAA 2280
2015 CAGCATTCGCGTGAAGATCTAGACCTTCAACAGAGCCAGATCCATCTACTGCTGGCCAA 2074
2281 TATGTTTCAAGTCCCTCAGGATTCAGTCTTGTGAAGTGAATCTTCTGGAAGTTGAG 2340
2075 TATGTTTCAAGTCCCTCAGGATTCAGTCTTGTGAAGTGAATCTTCTGGAAGTTGAG 2134
2341 TGTAACTCCGCTGATTTGACGGGCATATCCGTAGTTGGCAAGTGTGTTGTTGTTACGG 2400
2135 TGTAACTCCGCTGATTTGACGGGCATATCCGTAGTTGGCAAGTGTGTTGTTGTTACGG 2194
2401 AGAGTAACTCCAACTCTCTGAGAGTAGGACCAACCAACACACAGATCCAGCTGTT 2460
2195 AGAGTAACTCCAACTCTCTGAGAGTAGGACCAACCAACCAACACAGATCCAGCTGTT 2254
2461 GTACTTGTATCAATAGAAGACATTTCTCGATTTGACAGGATCAAGTGTTCAGSAGCGT 2520
2255 GTACTTGTATCAATAGAAGACATTTCTCGATTTGACAGGATCAAGTGTTCAGSAGCGT 2314
2521 ACTGATTTGACATTTCCAAAGCTCTGCTAGTGTGCAACCGATAGGTTGTAGAGTGTG 2580
2315 ACTGATTTGACATTTCCAAAGCTCTGCTAGTGTGCAACCGATAGGTTGTAGAGTGTG 2374
2581 CAATACATTTGGTACAATTTCAACCTTTGCAACTGACAGCTTGGTTGTGAAACAGCAT 2640
2375 CAATACATTTGGTACAATTTCAACCTTTGCAACTGACAGCTTGGTTGTGAAACAGCAT 2434
2641 CTTCAATTTGCGCAAGCTCCTTGTCTGATATCGAGAGCCAAACAGATACCTGGGAAT 2700
2435 CTTCAATTTGCGCAAGCTCCTTGTCTGATATCGAGAGCCAAACAGATACCTGGGAAT 2494
2701 CAATACCATTTGACGTTGAGACAGAGTCTGAGGCAACGAAATCTGATCAGGCTATT 2760
2495 CAATACCATTTGACGTTGAGACAGAGTCTGAGGCAACGAAATCTGATCAGGCTATT 2554
2761 TATCAGCAATAACTAGAACTTCTGTAAGGCCCCAGCAGGCAATGTCATATACTACAGGGGTG 2820
2555 TATCAGCAATAACTAGAACTTCTGTAAGGCCCCAGCAGGCAATGTCATATACTACAGGGGTG 2614
2821 ATGTGTCATTTTGAACCATCATCTTGGCAGCAGTAAACGAACTGTTTCTCGACCAATA 2880
2615 ATGTGTCATTTTGAACCATCATCTTGGCAGCAGTAAACGAACTGTTTCTCGACCAATA 2674
2881 TTTTGTCACTTTAGGAAACAGTTTCTGTTCCGTAAGCCATAGCAGTACTGCTGGGCGC 2940
2675 TTTTGTCACTTTAGGAAACAGTTTCTGTTCCGTAAGCCATAGCAGTACTGCTGGGCGC 2734
2941 CTCCTGCTAGCAGATACATTTAGCACCAACCTTGTGGGCAACGTAGATGACTTCTGGGG 3000
2735 CTCCTGCTAGCAGATACATTTAGCACCAACCTTGTGGGCAACGTAGATGACTTCTGGGG 2794
3001 TAAGGCTACCATCTTCTTAGTGGAGATGCAAAAAACAATTTCTTTGCAACAGCAACTT 3060

Db 2795 TAAGGGTACCATCTCTTAGGTGGAGATGCAAAAACAAATTTCTTTGCAACCGAACTT 2854
Qy 3061 TGGCAGGAACACCCAGCATCAGGAAGTGGAGGAGAGAAATGGCGTTCCACGAGGAATAT 3120
Db 2855 TGGCAGGAACACCCAGCATCAGGAAGTGGAGGAGAGAAATGGCGTTCCACGAGGAATAT 2914
Qy 3121 AGAGGCCAATCTTCTCAATAGGTCTTGCRAAACGAGAGGAGACTACACGAGGCGAAGTCT 3180
Db 2915 AGAGGCCAATCTTCTCAATAGGTCTTGCRAAACGAGAGGAGACTACACGAGGCGAAGTCT 2974
Qy 3181 CAATCTGCAACGCTCTCCGTTAGTTAGCTTCATGGAATTTCCGACGTTATCTATAGAGA 3240
Db 2975 CAATCTGCAACGCTCTCCGTTAGTTAGCTTCATGGAATTTCCGACGTTATCTATAGAGA 3034
Qy 3241 GATCAATGGCTCTCTTAACGTTATCTGGCAATTCGTAAGTTCCTCTGGGAAGGAGCTT 3300
Db 3035 GATCAATGGCTCTCTTAACGTTATCTGGCAATTCGTAAGTTCCTCTGGGAAGGAGCTT 3094
Qy 3301 CTAACACAGGTGTCTTCAAGCGACTCATCAAACTTTGGGAGTTAGTTCTTAAAGGGCTT 3360
Db 3095 CTAACACAGGTGTCTTCAAGCGACTCATCAAACTTTGGGAGTTAGTTCTTAAAGGGCTT 3154
Qy 3361 TGTCAACATTTGACGAACATPTGCGACAATTTGGTTTGGACTTAATTCOAATCTGTTCGG 3420
Db 3155 TGTCAACATTTGACGAACATPTGCGACAATTTGGTTTGGACTTAATTCOAATCTGTTCGG 3214
Qy 3421 TTTTCTGATAGGACGAGGAGGCACTTCAATTTCTTGTGGAGGCGCTTGAACAGCT 3480
Db 3215 TTTTCTGATAGGACGAGGAGGCACTTCAATTTCTTGTGGAGGCGCTTGAACAGCT 3274
Qy 3481 CAATTTTGCACAATTCATACGACCTTCAGAAGGAGCTTCTTTAGGTTTGGATTTCTTCT 3540
Db 3275 CAATTTTGCACAATTCATACGACCTTCAGAAGGAGCTTCTTTAGGTTTGGATTTCTTCT 3334
Qy 3541 TAGTTGTTCTTGGTGTATCTGCTTCCACCTCGTCCAACTGTCACCGTACTTGGCACTATA 3600
Db 3335 TAGTTGTTCTTGGTGTATCTGCTTCCACCTCGTCCAACTGTCACCGTACTTGGCACTATA 3394
Qy 3601 ACTTCATATCAGGTTCTCTCCACCTCGTCCAACTGTCACCGTACTTGGCACTATA 3660
Db 3395 ACTTCATATCAGGTTCTCTCCACCTCGTCCAACTGTCACCGTACTTGGCACTATA 3454
Qy 3661 CTAATGCAAAATAAAATAAGTCAGACATTCGCCAGGCTATATCTTCCCTTGGATTTAGCTT 3720
Db 3455 CTAATGCAAAATAAAATAAGTCAGACATTCGCCAGGCTATATCTTCCCTTGGATTTAGCTT 3514
Qy 3721 CTGCAAGTTCAATCAGCTCTCCCTCAATTTTAGGCTTCAACAACTTCGTCGTAATA 3780
Db 3515 CTGCAAGTTCAATCAGCTCTCCCTCAATTTTAGGCTTCAACAACTTCGTCGTAATA 3574
Qy 3781 ACCGTTTGGTATAGAAACCTTCTCGAGCAATTCCTTACGATCCCAAGGTGGCTTCCA 3840
Db 3575 ACCGTTTGGTATAGAAACCTTCTCGAGCAATTCCTTACGATCCCAAGGTGGCTTCCA 3634
Qy 3841 TGGCTCTAAGACCCCTTTGATTTGGCAAAACAGGAAGTGGCTTCCAGTGCAGAAACCAA 3900
Db 3635 TGGCTCTAAGACCCCTTTGATTTGGCAAAACAGGAAGTGGCTTCCAGTGCAGAAACCAA 3694
Qy 3901 CACCTGTTGTTCAACCAAAATTTCAAGAGTCTCCATCACAATCCAATTCGATACCCA 3960
Db 3695 CACCTGTTGTTCAACCAAAATTTCAAGAGTCTCCATCACAATCCAATTCGATACCCA 3754
Qy 3961 GCACTTTTGGTTGCTCCAGATGACACTTTTATACCAAAACCGGTGACGAGATT 4020
Db 3755 GCACTTTTGGTTGCTCCAGATGACACTTTTATACCAAAACCGGTGACGAGATT 3814
Qy 4021 GGTAGACTCCAGTTTGTCTCTTATAGCTTCGGGAATAGACTTTTGGAGAGGTACACCA 4080
Db 3815 GGTAGACTCCAGTTTGTCTCTTATAGCTTCGGGAATAGACTTTTGGAGAGGTACACCA 3874
Qy 4081 GGCCCAACAGTAATTAGAAGAGTCAGCCACCAAAAGTAGTGAATAGACCAATCGGGCGGT 4140
Db 3875 GGCCCAACAGTAATTAGAAGAGTCAGCCACCAAAAGTAGTGAATAGACCAATCGGGCGGT 3934

Qy 4141 CAGTAGTCAAGACGCCCAACAAAAATTTCACTGACAGGGAACTTTTGTGACATCTTCAGAAA 4200
Db 3935 CAGTAGTCAAGACGCCCAACAAAAATTTCACTGACAGGGAACTTTTGTGACATCTTCAGAAA 3994
Qy 4201 GTTCGTATTTCAGTGTCAATTCGCCGAGCATCAATTAATGGGATTTATACAGAAAGCAACAG 4260
Db 3995 GTTCGTATTTCAGTGTCAATTCGCCGAGCATCAATTAATGGGATTTATACAGAAAGCAACAG 4054
Qy 4261 TGAAGTGCATCTCTACCAACTTTTGGGTCTCAGAAAAAGCATATAACAGTTCTACTACGC 4320
Db 4055 TGAAGTGCATCTCTACCAACTTTTGGGTCTCAGAAAAAGCATATAACAGTTCTACTACGC 4114
Qy 4321 CATTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAAAGGACAGCATCTAGCAT 4380
Db 4115 CATTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAAAGGACAGCATCTAGCAT 4174
Qy 4381 TAGCGGGCAAGGATGCACTTTTATCAACAGGGTCTCTATAGATAACCCCTAGCGCTGGGA 4440
Db 4175 TAGCGGGCAAGGATGCACTTTTATCAACAGGGTCTCTATAGATAACCCCTAGCGCTGGGA 4234
Qy 4441 TCATCTTTGGACAACCTTTCTTCGCAAACTAGSTCCAAAATCACTTCATTGATACCAT 4500
Db 4235 TCATCTTTGGACAACCTTTCTTCGCAAACTAGSTCCAAAATCACTTCATTGATACCAT 4294
Qy 4501 TATTGTACAACTTGAGCAAGTTTGTGATCAGCTCTCAAAATTTGGTCTCTGTAAACGATG 4560
Db 4295 TATTGTACAACTTGAGCAAGTTTGTGATCAGCTCTCAAAATTTGGTCTCTGTAAACGATG 4354
Qy 4561 ACTCAACTTTGCACTTAATTTGAAGTCACTAGTGCATTTAGTGAACTTTGATTCAGGTTGTGA 4620
Db 4355 ACTCAACTTTGCACTTAATTTGAAGTCACTAGTGCATTTAGTGAACTTTGATTCAGGTTGTGA 4414
Qy 4621 GCTGGTCAAGCAGCATAGGGAACACGGCTTTTCTTACCAAACTCAAGGAATTTATCAAACT 4680
Db 4415 GCTGGTCAAGCAGCATAGGGAACACGGCTTTTCTTACCAAACTCAAGGAATTTATCAAACT 4474
Qy 4681 CTGCAACACTTTCGATGCAAGTGTAGTCAAGGAAATGTCTATCTTGAAGTCGACAGTGAG 4740
Db 4475 CTGCAACACTTTCGATGCAAGTGTAGTCAAGGAAATGTCTATCTTGAAGTCGACAGTGAG 4534
Qy 4741 TGTAGTCTTGAGAAATTTGAAAGCGTATTTTATATCAGTGCATTCAGTGCATCAGGAGA 4800
Db 4535 TGTAGTCTTGAGAAATTTGAAAGCGTATTTTATATCAGTGCATTCAGTGCATCAGGAGA 4594
Qy 4801 TCTCTTACCCCGAGCGCATCTGTGGCCGA----- 4828
Db 4595 TCTCTTACCCCGAGCGCATCTGTGGCCGA----- 4654
Qy 4829 ----- 4828
Db 4655 CTCTGAGAGAGGTGTGTGCTACTATACAGGCGCTGAATCGCCCATCATCCAGCCAGA 4714
Qy 4829 ----- 4828
Db 4715 AAGTGAGGAGCCACGGTTGATGAGAGCTTTGTTGTAGGTGGAACCAAGTTGGTGAATTGGA 4774
Qy 4829 ----- 4828
Db 4775 ACTTTTGTCTTCCAGCGAAACGGTCTGCTGTCTGCGGAAGATGCGTGATCTGATCTCTCA 4834
Qy 4829 ----- 4828
Db 4835 ACTCAGAAAAAGTTCGATTTATTTCACAAAGCGCCGCTCCCGTCAAGTCAGCGTAATGCT 4894
Qy 4829 ----- 4828
Db 4895 CTGCCAGTGTACAAACCAATTAACCAATTTCTGATTAGAAAAAACTCATCGAGCATCAAAATG 4954
Qy 4829 ----- 4828
Db 4955 AAACGTCAATTTATTATATCATATCAGGATTTCAATACCATATTTTGTAAAAAGCGCTTCTG 5014

QY 4829 ----- 4828
Db 5015 TAATGAAGAGAGAAAACTCACCGAGGCGTTCATAGATGGCAAGATCCTGGTATCGGTC 5074
QY 4829 ----- 4828
Db 5075 TGGATTCCGACTCGTCAACATCAATACAACTTAATTTCCCTCGTCAAAAATAAG 5134
QY 4829 ----- 4828
Db 5135 GTTATCAAGTGAGAAATCAACATAGTGAACGATCCGGTGAGATGGCAAAAGCTT 5194
QY 4829 ----- 4828
Db 5195 ATGCATTTCTTCAGACTTGTTCACAGCGCAGCCATTACGTCGTTCATCAAAATCACT 5254
QY 4829 ----- 4828
Db 5255 CGCATCAACCAACCGTTATTTCATTCGTGATTCGCTGAGAGAGAAATACGGATC 5314
QY 4829 ----- 4828
Db 5315 GCTGTTAAAGGACAATTACAAACAGAAATCGAATGCAACGGCGCAGGAACACTGCCAG 5374
QY 4829 ----- 4828
Db 5375 CGCATCAACAAATTTTTCACCTGAATCAGGATATTCTTAAATACCTGGAATGCTGTTT 5434
QY 4829 ----- 4828
Db 5435 CCCGGGATCCAGTGGTGAGTAACCATGATCATCAGGAGTACGATAAATGCTTGAT 5494
QY 4829 ----- 4828
Db 5495 GGTCCGAAGAGCATAAATCCGTCAGCAGCTTTAGTCTGACCATCTCATCTGTAACATC 5554
QY 4829 ----- 4828
Db 5555 ATTGGCAACGCTACTTTTGGCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCAT 5614
QY 4829 ----- 4828
Db 5615 CAATCGATAGATTGTCACCTGATTGCCGACATTATCGCGACATTATCGAGCCCATTTATACCCATA 5674
QY 4829 ----- 4828
Db 5675 TAAATCAGCATCCATGTTGGAAATTAATTCGGGCTCGAGCAAGCGTTTCCCGTTGAAT 5734
QY 4829 ----- 4828
Db 5735 ATGGCTATAACACCCCTGTTACTGTTTACTGTTTATGTAAGCAGACAGTTTATTGTTATGA 5794
QY 4829 ----- 4828
Db 5795 TGATATATTTTATCTTGTGCAATGTAACATCAGAGATTTTGACACAAACGCTGCTTTC 5854
QY 4829 ----- CTTGAGGTCCGATCACCGGCCACAGGTGCGGTCTCGCGCCCTATAT 4879
Db 5855 CCCCCCCCCCTGACAGTTCGATCACCGGCCACAGGTGCGGTCTCGCGCCCTATAT 5914
QY 4880 CGCGACATCAACGATGGGAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTGTTT 4939
Db 5915 CGCGACATCAACGATGGGAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTGTTT 5974
QY 4940 CGCGTGGGTATGTTGGAGCCCGCTGCGCGGAGCTGTTGGCGCATCTCCCTGCA 4999
Db 5975 CGCGTGGGTATGTTGGAGCCCGCTGCGCGGAGCTGTTGGCGCATCTCCCTGCA 6034
QY 5000 TGCACCATCTCTCGCGCGGCTGCTCAACGGGCTCAACCTACTACTGCGGCTGCTTCCCT 5059
Db 6035 TGCACCATCTCTCGCGCGGCTGCTCAACGGGCTCAACCTACTACTGCGGCTGCTTCCCT 6094
QY 5060 AATGAGGAGTCCGATTAAGGAGAGCGCTCGAGTATCTATGATGGAAGTATGGAAATGGT 5119

Db 6095 AATGAGGAGTCCGATTAAGGAGAGCGCTCGAGTATCTATGATGGAAGTATGGAAATGGT 6154
QY 5120 GATACCCGATCTCTCAGTGTCTTGAAGTCTTCCATCAGATTATGCCCAACTAAAGCAAC 5179
Db 6155 GATACCCGATCTCTCAGTGTCTTGAAGTCTTCCATCAGATTATGCCCAACTAAAGCAAC 6214
QY 5180 CGGAGGAGGAGATTTATGTTAAATTTCTCTGACTTTTGGTTCATCAGTACGACTGCAACTG 5239
Db 6215 CGGAGGAGGAGATTTATGTTAAATTTCTCTGACTTTTGGTTCATCAGTACGACTGCAACTG 6274
QY 5240 TGAGACTATCTCGGTATGACAGCAGAAATGCTCTTCTGGAGACAGTAAATGAAGTCCC 5299
Db 6275 TGAGACTATCTCGGTATGACAGCAGAAATGCTCTTCTGGAGACAGTAAATGAAGTCCC 6334
QY 5300 ACCAATAAGAAATCTCTTGTATCAGGAACAAATCTTCTTGTTCGAACTTTTTCGGTGCC 5359
Db 6335 ACCAATAAGAAATCTCTTGTATCAGGAACAAATCTTCTTGTTCGAACTTTTTCGGTGCC 6394
QY 5360 TTGAACATAAATGTAGAGTGGATATGTCGGGTAGGAATGGAGCGGCAATGCTTACC 5419
Db 6395 TTGAACATAAATGTAGAGTGGATATGTCGGGTAGGAATGGAGCGGCAATGCTTACC 6454
QY 5420 TTCTGGACCTTCAAGAGTATGAGGTTTGTAGATCTATGATGCCAACTTCAGTGACAA 5479
Db 6455 TTCTGGACCTTCAAGAGTATGAGGTTTGTAGATCTATGATGCCAACTTCAGTGACAA 6514
QY 5480 GTTGCTATTTCTGTTCAACCAATCCGAATCCAGAGAAATCAAGTGTGTTGTCTACTATT 5539
Db 6515 GTTGCTATTTCTGTTCAACCAATCCGAATCCAGAGAAATCAAGTGTGTTGTCTACTATT 6574
QY 5540 GATCCAGCAGTCCGGTCTTGAACCTGACAAATAGTGTCTCGTGTGTTTGAAGTCACTTT 5599
Db 6575 GATCCAGCAGTCCGGTCTTGAACCTGACAAATAGTGTCTCGTGTGTTTGAAGTCACTTT 6634
QY 5600 TGTATGAATAAATCTAGTCTTTGATCTAAATAATCTTGAAGCAAGGCAATAAATACC 5659
Db 6635 TGTATGAATAAATCTAGTCTTTGATCTAAATAATCTTGAAGCAAGGCAATAAATACC 6694
QY 5660 CAATCTAAATCTTTTAAACGTTAAAGGCAAGTATGTCCTGCTGTTTAAAGTCACTTT 5719
Db 6695 CAATCTAAATCTTTTAAACGTTAAAGGCAAGTATGTCCTGCTGTTTAAAGTCACTTT 6754
QY 5720 AAATCAGCTCGTAGTCTGATCTCATCAACTTGAAGGCACTATCTTGTGTTTGAAGAAAT 5779
Db 6755 AAATCAGCTCGTAGTCTGATCTCATCAACTTGAAGGCACTATCTTGTGTTTGAAGAAAT 6814
QY 5780 TTGCGGAGATCGGATATCGAGAAAGGTAACGTTAAAGTAAAGTAAATTTATCTCAA 5839
Db 6815 TTGCGGAGATCGGATATCGAGAAAGGTAACGTTAAAGTAAAGTAAATTTATCTCAA 6874
QY 5840 GATCTCTGCTCGCGCTTTCGGTGATGACGTTGAAACCTCTGACACATCGAGTCCCG 5899
Db 6875 GATCTCTGCTCGCGCTTTCGGTGATGACGTTGAAACCTCTGACACATCGAGTCCCG 6934
QY 5900 GAGACGGTCAACGCTTGTCTGTAAGCGGATCCCGGAGCAGACAAAGCCGTCAGGGGCG 5959
Db 6935 GAGACGGTCAACGCTTGTCTGTAAGCGGATCCCGGAGCAGACAAAGCCGTCAGGGGCG 6994
QY 5960 TCAGCGGTGTTGGCGGTGTCGGGCGCAGCATGACCCAGTCACTGAGGATAGGGGA 6019
Db 6995 TCAGCGGTGTTGGCGGTGTCGGGCGCAGCATGACCCAGTCACTGAGGATAGGGGA 7054
QY 6020 GTGTATCTGCTTAACTATGCGGCATCAGAGCAGATTTGTAAGAGTGCACCATATGC 6079
Db 7055 GTGTATCTGCTTAACTATGCGGCATCAGAGCAGATTTGTAAGAGTGCACCATATGC 7114
QY 6080 GTGTATCTGCTTAACTATGCGGCATCAGAGCAGATTTGTAAGAGTGCACCATATGC 6139
Db 7115 GTGTATCTGCTTAACTATGCGGCATCAGAGCAGATTTGTAAGAGTGCACCATATGC 7174
QY 6140 CTTCTGCTCACTGACTCGCTGCGCTCGGTCTGCGTCTGCGGCGAGCGGTATCAGTCACT 6199

Db 7175 CCTCGCTCACTGACTCGCTCGGCTCGGTCGTTCCGCTCGGCGAGCGGTATCAGCTCACT 7234
Qy 6200 CAAAGCGCGTAAATACGGTTATCCACAGAAATCAGGGGATAACCGCAGGAAGAACATGTCAG 6259
Db 7235 CAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACCGCAGGAAGAACATGTCAG 7294
Qy 6260 CAAAGCGGACGAAAGCCAGGAAACCTGTAAGGCGCGTTGCTGCGCTTTTCCATA 6319
Db 7295 CAAAGGCGCAGCAAGGCGCAGGAAACCTGTAAGGCGCGTTGCTGCGCTTTTCCATA 7354
Qy 6320 GGCCTCGGCCCCCTGACGAGCATCACAAATCGACGCTCAAGTCAGAGGTGGCGAAACC 6379
Db 7355 GGCCTCGGCCCCCTGACGAGCATCACAAATCGACGCTCAAGTCAGAGGTGGCGAAACC 7414
Qy 6380 CGACAGGACTATAAGATACCAAGCGGTTCCCGCTGGAAGCTCCCTCGTGCGCTCTCCTG 6439
Db 7415 CGACAGGACTATAAGATACCAAGCGGTTCCCGCTGGAAGCTCCCTCGTGCGCTCTCCTG 7474
Qy 6440 TTCCGACCGCTGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGGCG 6499
Db 7475 TTCCGACCGCTGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGGCG 7534
Qy 6500 TTCTCTCATAGCTCAGCGTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAAGCTGG 6559
Db 7535 TTCTCTCATAGCTCAGCGTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAAGCTGG 7594
Qy 6560 GCTGTGTGCAAGAACCCCGCTGACGCGGACCGCTGCGCTTATCCGTTAACTATCGTC 6619
Db 7595 GCTGTGTGCAAGAACCCCGCTGACGCGGACCGCTGCGCTTATCCGTTAACTATCGTC 7654
Qy 6620 TTGAGTCCAAACCCGTAAGACACGACTTATCGCACCTGGCAGCGCCACTGGTAACAGGA 6679
Db 7655 TTGAGTCCAAACCCGTAAGACACGACTTATCGCACCTGGCAGCGCCACTGGTAACAGGA 7714
Qy 6680 TTAGCAGAGCGAGTATAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCCTAACTACG 6739
Db 7715 TTAGCAGAGCGAGTATAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCCTAACTACG 7774
Qy 6740 GCTACACTAGAGACAGTATTTGGTATCTCGCTCTGCTGACGCGGATTAACCTTCGAA 6799
Db 7775 GCTACACTAGAGACAGTATTTGGTATCTCGCTCTGCTGACGCGGATTAACCTTCGAA 7834
Qy 6800 AAAGAGTTGGTAGCTCTTGATCCGGCAAAACCAACCGCTGGTAGCGGTGTTTGTG 6859
Db 7835 AAAGAGTTGGTAGCTCTTGATCCGGCAAAACCAACCGCTGGTAGCGGTGTTTGTG 7894
Qy 6860 TTGCAAGCAGCAGATACGGCGGAAAGAAAGATCTCAAGAGATCTTGTGATCTTT 6919
Db 7895 TTGCAAGCAGCAGATACGGCGGAAAGAAAGATCTCAAGAGATCTTGTGATCTTT 7954
Qy 6920 CTACGGGCTCTGAGCTCAGTGGAAACGAAACTCAGGTTAAGGATTTGCTCATGAGAT 6979
Db 7955 CTACGGGCTCTGAGCTCAGTGGAAACGAAACTCAGGTTAAGGATTTGCTCATGAGAT 8014
Qy 6980 TATCAAAAGGATCTTCACTTAGATCTCTTTTAAATTAAGTGTAAATCAATCT 7039
Db 8015 TATCAAAAGGATCTTCACTTAGATCTCTTTTAAATTAAGTGTAAATCAATCT 8074
Qy 7040 AAAGTATATAGTAAATCTGCTGACGTTACCAATGCTTAATCAGTGAGGACCTTA 7099
Db 8075 AAAGTATATAGTAAATCTGCTGACGTTACCAATGCTTAATCAGTGAGGACCTTA 8134
Qy 7100 TCTCAGCGATCTGCTATTTGCTTATCCATGATGCTGACTCCCGCTGGTAGATAA 7159
Db 8135 TCTCAGCGATCTGCTATTTGCTTATCCATGATGCTGACTCCCGCTGGTAGATAA 8194
Qy 7160 CTACGATACGGAGGCTTACATCTGGCCCGAGTGTGCAATGATACCGGAGACCCAC 7219
Db 8195 CTACGATACGGAGGCTTACATCTGGCCCGAGTGTGCAATGATACCGGAGACCCAC 8254
Qy 7220 GCTCACCGGCTCCAGATTTATCAGCAATACACGAGCGGAGCGGAGCGGAGGAG 7279
Db 8255 GCTCACCGGCTCCAGATTTATCAGCAATACACGAGCGGAGCGGAGCGGAGGAGGAG 8314

RESULT 5
US-10-112-221A-11
; Sequence 11, Application US/10112221A

Qy 7280 GTGGTCTCTGCAACTTTTATCCGCTCCATCCAGTCTAATTAATTTGTCGCGGAGCTTAGAG 7339
Db 8315 GTGGTCTCTGCAACTTTTATCCGCTCCATCCAGTCTAATTAATTTGTCGCGGAGCTTAGAG 8374
Qy 7340 TAAAGTAGTTCGCGAGTTAATAGTTTCGCAACGTTGTCGAGGCAATCGTGG 7399
Db 8375 TAAAGTAGTTCGCGAGTTAATAGTTTCGCAACGTTGTCGAGGCAATCGTGG 8434
Qy 7400 TGTACGCTCGTGGTTGGTATGGCTTCATTCAGCTCCGGTTCGCTCCGATCGTTG 7459
Db 8435 TGTACGCTCGTGGTTGGTATGGCTTCATTCAGCTCCGGTTCGCTCCGATCGTTG 8494
Qy 7460 TTAATGATCCCGCACTGTTGTGCAAAAGAGCGGTAGCTCCTTCGGTTCCTCGATCGTTG 7519
Db 8495 TTAATGATCCCGCACTGTTGTGCAAAAGAGCGGTAGCTCCTTCGGTTCCTCGATCGTTG 8554
Qy 7520 TCAGAGTAAGTTGGCGCGAGTGTATCACTCATGTTATGGAGACTGCAATTTCTC 7579
Db 8555 TCAGAGTAAGTTGGCGCGAGTGTATCACTCATGTTATGGAGACTGCAATTTCTC 8614
Qy 7580 TTAATGATCCCGCACTGTTGTGCAAAAGAGCGGTAGCTCCTTCGGTTCCTCGATCGTTG 7639
Db 8615 TTAATGATCCCGCACTGTTGTGCAAAAGAGCGGTAGCTCCTTCGGTTCCTCGATCGTTG 8674
Qy 7640 TCTGAGATAGTGTATGCGCGGACCGAGTTGCTTTGCCGCGCTCAACACCGGATAATA 7699
Db 8675 TCTGAGATAGTGTATGCGCGGACCGAGTTGCTTTGCCGCGCTCAACACCGGATAATA 8734
Qy 7700 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCTGGAAGAGCTTCTTCGGGCGAA 7759
Db 8735 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCTGGAAGAGCTTCTTCGGGCGAA 8794
Qy 7760 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCGATGTAACCCACTCTGTCACCCA 7819
Db 8795 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCGATGTAACCCACTCTGTCACCCA 8854
Qy 7820 ACTGATCTTCAGCATCTTTTACCTTCCAGCGGTTCTGGTGAGCAAAACAGGAGGC 7879
Db 8855 ACTGATCTTCAGCATCTTTTACCTTCCAGCGGTTCTGGTGAGCAAAACAGGAGGC 8914
Qy 7880 AAAAAATCCGCAAAAGGGAATAAGGCGGACACGGAATGTTGAATCTCATCTCTTCC 7939
Db 8915 AAAAAATCCGCAAAAGGGAATAAGGCGGACACGGAATGTTGAATCTCATCTCTTCC 8974
Qy 7940 TTTTTCATATTTTGAAGCATTTATCAGGTTTATGCTCTCATGAGCGGATACATTTG 7999
Db 8975 TTTTTCATATTTTGAAGCATTTATCAGGTTTATGCTCTCATGAGCGGATACATTTG 9034
Qy 8000 AATGTAATTTAGAAAAATAAACAATAAGGTTCCGCGACATTTCCCGAAAAAGTCCAC 8059
Db 9035 AATGTAATTTAGAAAAATAAACAATAAGGTTCCGCGACATTTCCCGAAAAAGTCCAC 9094
Qy 8060 CTGACGCTTAAGAAACCATTTATCATGACATTAACCTTAAATAAGGCGTATCACGA 8119
Db 9095 CTGACGCTTAAGAAACCATTTATCATGACATTAACCTTAAATAAGGCGTATCACGA 9154
Qy 8120 GGCCTTTTCGCTTCAAGAAATTAATCTCATGTTTGAAGCTTATCATGATAGCTGAC 8179
Db 9155 GGCCTTTTCGCTTCAAGAAATTAATCTCATGTTTGAAGCTTATCATGATAGCTGAC 9214
Qy 8180 TCATGTTGTTATTTGTAATAGACGAGATCGGGAACACTGAAAAATACAGTTATTAT 8239
Db 9215 TCATGTTGTTATTTGTAATAGACGAGATCGGGAACACTGAAAAATACAGTTATTAT 9274
Qy 8240 CG 8241
Db 9275 CG 9276


```

; Publication No. US20030166851A1
; GENERAL INFORMATION:
; APPLICANT: Madison, Edwin
; APPLICANT: Ong, Edgar O.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING A TRANSMEMBRANE SERINE PROTEASE
; FILE REFERENCE: 24745-1615
; CURRENT APPLICATION NUMBER: US/10/112,221A
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: 60/279,228
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/291,501
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 9276
; TYPE: DNA
; ORGANISM: Pichia pastoris
US-10-112-221A-11

Query Match      79.6%; Score 6561.6; DB 15; Length 9276;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;

QY      1 AGATCTAACATCCAAAGACGAAAGGTGGAATGAAACCTTTTGGCATCCCGACATCCACAG 60
DB      |||
QY      1 AGATCTAACATCCAAAGACGAAAGGTGGAATGAAACCTTTTGGCATCCCGACATCCACAG 60
DB      |||
QY      61 GTCCATCTCACACATAAGTGCACAAAGCAACAGAGGGGATACACTAGCAGACAGCGT 120
DB      |||
QY      61 GTCCATCTCACACATAAGTGCACAAAGCAACAGAGGGGATACACTAGCAGACAGCGT 120
DB      |||
QY      121 TGCAACAGCAGACCTCCACTCTCTCTCTCTCAACACCCACTTTTGCATCCGAAACACC 180
DB      |||
QY      121 TGCAACAGCAGACCTCCACTCTCTCTCTCTCAACACCCACTTTTGCATCCGAAACACC 180
DB      |||
QY      181 AGCCAGATTATGGGCTTGANTGGAGCTCGTCTCAATCCAAATCCCTCTATAGGCTACTA 240
DB      |||
QY      181 AGCCAGATTATGGGCTTGANTGGAGCTCGTCTCAATCCAAATCCCTCTATAGGCTACTA 240
DB      |||
QY      241 ACACCATGACTTTATAGCTGTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
DB      |||
QY      241 ACACCATGACTTTATAGCTGTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
DB      |||
QY      301 TTTCGAAATGCAACAGCTCCGCAATACACCCGAACATCACTCCAGATGAGGGCTTTCTG 360
DB      |||
QY      301 TTTCGAAATGCAACAGCTCCGCAATACACCCGAACATCACTCCAGATGAGGGCTTTCTG 360
DB      |||
QY      361 AGTGTGGGTCAAAATAGTTTCAATGTTCCCAAAAGGGCCCAAACTGACAGCTTTAAAGCT 420
DB      |||
QY      361 AGTGTGGGTCAAAATAGTTTCAATGTTCCCAAAAGGGCCCAAACTGACAGCTTTAAAGCT 420
DB      |||
QY      421 GTCTTGGAACTTAATATGACAAAAGCGTGATCTCATCCAAAGTGAACCTAAAGTTTGGTTG 480
DB      |||
QY      481 TTGAAATGCTAACGCGCTAGTTGTTCAAAAGAAACTTCCAAAGTCCGCATACCGTTGT 540
DB      |||
QY      481 TTGAAATGCTAACGCGCTAGTTGTTCAAAAGAAACTTCCAAAGTCCGCATACCGTTGT 540
DB      |||
QY      541 CTGTGTTGGTATGATGAGAAATGCTCAAAATAATCTCATTAATGCTTAGCGCAGTCT 600
DB      |||
QY      541 CTGTGTTGGTATGATGAGAAATGCTCAAAATAATCTCATTAATGCTTAGCGCAGTCT 600
DB      |||
QY      601 CTCATCGCTTCTGAACCCCGGTGACCTGTGCGCAACCAATGCGGAAACACCCCGCT 660
DB      |||
QY      601 CTCATCGCTTCTGAACCCCGGTGACCTGTGCGCAACCAATGCGGAAACACCCCGCT 660
DB      |||
QY      661 TTTTGGATGATATGATGCTTCCACATGATGCTTCCAAAGTCTGCGGAAATACT 720
DB      |||
QY      661 TTTTGGATGATATGATGCTTCCACATGATGCTTCCAAAGTCTGCGGAAATACT 720
DB      |||

721 GCTGATAGCTTAAACGTTTCATGATCAAAATTTAACTGTTCTAAACCCCTACTTGACAGCAAT 780
721 GCTGATAGCTTAAACGTTTCATGATCAAAATTTAACTGTTCTAAACCCCTACTTGACAGCAAT 780
781 ATATATAACAGAGAGAGCTGCCCTGCTCTTAAACCTTTTATCATCATATTAAGCTT 840
781 ATATATAACAGAGAGAGCTGCCCTGCTCTTAAACCTTTTATCATCATATTAAGCTT 840
841 ACTTTTCAATTTGCGACTGTTTCCAAATTTGACAGCTTTTGAATTTAAACGACTTTTAAACGA 900
841 ACTTTTCAATTTGCGACTGTTTCCAAATTTGACAGCTTTTGAATTTAAACGACTTTTAAACGA 900
901 CAACTTGAGAGATCAAAAACAACTAATTTATTCGAAGGATCCAAAGCATGAGATTTCT 960
901 CAACTTGAGAGATCAAAAACAACTAATTTATTCGAAGGATCCAAAGCATGAGATTTCT 960
961 TCAATTTTACTGCACTTTTATTCGCAAGCATCTCCGCAATCTCTCCAGTCAACACT 1020
961 TCAATTTTACTGCACTTTTATTCGCAAGCATCTCTCCGCAATCTCTCCAGTCAACACT 1020
1021 ACAACAGAGATGAAACGGCAAAATTCGGCTGAGAGCTGTCTATCGGTTACTCAGATTTA 1080
1021 ACAACAGAGATGAAACGGCAAAATTCGGCTGAGAGCTGTCTATCGGTTACTCAGATTTA 1080
1081 GAAAGGGATTTGCGATGTTGTTTCCCAATTTTCCAAACAGCAAAATTAACGGGTTATTG 1140
1081 GAAAGGGATTTGCGATGTTGTTTCCCAATTTTCCAAACAGCAAAATTAACGGGTTATTG 1140
1141 TTTTAAATACTACTATTTGCCAGCATTTGCTGCTTAAAGAGAGGGGTATCTCTCGAGAA 1200
1141 TTTTAAATACTACTATTTGCCAGCATTTGCTGCTTAAAGAGAGGGGTATCTCTCGAGAA 1200
1201 AGAGAGGCTGAAGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAG 1260
1201 AGAGAGGCTGAAGCTTAAATTTGGCTTCCCGGTGTCAACGCTTCCAGCAGTGCAGGAG 1260
1217 GCTTAAATGCGGTAGTTTATATCAAGAGAAATGTCCTTCTTAACTAGTGGGCTAGAAATTCCTTAGG 1440
1217 GCTTAAATGCGGTAGTTTATATCAAGAGAAATGTCCTTCTTAACTAGTGGGCTAGAAATTCCTTAGG 1440
1441 CGGCGCGGAATTAATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTGGGCACTTAC 1500
1441 CGGCGCGGAATTAATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTGGGCACTTAC 1500
1501 AGAAGACCGGCTTTGCTAGATTTCTAATCAAGAGGATGTCAGAAATGCCATTTGCTGAGAG 1560
1501 AGAAGACCGGCTTTGCTAGATTTCTAATCAAGAGGATGTCAGAAATGCCATTTGCTGAGAG 1560
1561 ATGCAAGCTTCAATTTTGTATCTTTTATTTTCTTAACTATATAGTATAGGATTTTCT 1620
1561 ATGCAAGCTTCAATTTTGTATCTTTTATTTTCTTAACTATATAGTATAGGATTTTCT 1620
1621 GTCAATTTGTTCTCTCGTACGAGCTGCTCTGATCAGCCTATCTGATGAGGATTTTCT 1740
1621 GTCAATTTGTTCTCTCGTACGAGCTGCTCTGATCAGCCTATCTGATGAGGATTTTCT 1740
1741 ACTCCTCTTCCAGAGTACAGAAATTAAGTGAAGGTTCTGTTGCGAGCTTATCGATTA 1800
1741 ACTCCTCTTCCAGAGTACAGAAATTAAGTGAAGGTTCTGTTGCGAGCTTATCGATTA 1800
1800 ACTCCTCTTCCAGAGTACAGAAATTAAGTGAAGGTTCTGTTGCGAGCTTATCGATTA 1800
1800 ACTCCTCTTCCAGAGTACAGAAATTAAGTGAAGGTTCTGTTGCGAGCTTATCGATTA 1800
1860 GCTTAAATGCGGTAGTTTATATCAAGAGAAATGCTTAAACGCTTCCAGGACCGCTGATGAA 1860

```


Db 1595 GCTTTATGCGGTAGTTTATCAAGTTAAATGCTAACGCGTACGACCGGTGATGAA 1654
Qy 1861 ATCTAAAGTGGCTCATGCTCATCTCGGCAACCGTACCCCTGGATGCTGTAGGATAGG 1920
Db 1655 ATCTAAAGTGGCTCATGCTCATCTCGGCAACCGTACCCCTGGATGCTGTAGGATAGG 1714
Qy 1921 CTTGGTATGCGGTACTGCGGGGCTCTTGGGGATATCGTCCATTCGCAAGCATCGC 1980
Db 1715 CTTGGTATGCGGTACTGCGGGGCTCTTGGGGATATCGTCCATTCGCAAGCATCGC 1774
Qy 1981 CAGTCACATAGCGGTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGACCCGT 2040
Db 1775 CAGTCACATAGCGGTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGACCCGT 1834
Qy 2041 TCTCGGAGCTGTCGACCGCTTTGGCGCGGCCAGTCTGCTCGTCTGCTACTTGG 2100
Db 1835 TCTCGGAGCTGTCGACCGCTTTGGCGCGGCCAGTCTGCTCGTCTGCTACTTGG 1894
Qy 2101 AGCCACTATCGACTACGGGATCATGGCGACACACCCGCTCTGTGGATCTATCGAATCTA 2160
Db 1895 AGCCACTATCGACTACGGGATCATGGCGACACACCCGCTCTGTGGATCTATCGAATCTA 1954
Qy 2161 ATGTAAAGTTAAATCTCTAAATTAATTAATTAAGTCCAGTTCTCCATAGCAACCTTAA 2220
Db 1955 AATGTAAGTTAAATCTCTAAATTAATTAATTAAGTCCAGTTCTCCATAGCAACCTTAA 2014
Qy 2221 CAGCAATTCGGGTAGCATCTAGACCTTCAACAGCAGCCAGATCCATCACTCTTGGCCAA 2280
Db 2015 CAGCAATTCGGGTAGCATCTAGACCTTCAACAGCAGCCAGATCCATCACTCTTGGCCAA 2074
Qy 2281 TATGTTTCAGTCCCTCAGGAGTTAGTCTTGTGAAGTGAATCTTCTGGAAGTTGCG 2340
Db 2075 TATGTTTCAGTCCCTCAGGAGTTAGTCTTGTGAAGTGAATCTTCTGGAAGTTGCG 2134
Qy 2341 TGTTAACTCCCTGTATTGAAGGCGATATCGTAGCTTGGCAAGTGGTGTAGCGG 2400
Db 2135 TGTTAACTCCCTGTATTGAAGGCGATATCGTAGCTTGGCAAGTGGTGTAGCGG 2194
Qy 2401 AGGAGTAATCTCCAACTCTCTGAGAGTAGGACCAACAAACACAGATCCACGCTGTT 2460
Db 2195 AGGAGTAATCTCCAACTCTCTGAGAGTAGGACCAACAAACACAGATCCACGCTGTT 2254
Qy 2461 GTACTTGATCAACATAGAAGAGATTTCTGATTTGCGAGATCAAGTTTTCAGGAGCGT 2520
Db 2255 GTACTTGATCAACATAGAAGAGATTTCTGATTTGCGAGATCAAGTTTTCAGGAGCGT 2314
Qy 2521 ACTGATTCGACATTTCCAAAGCCTGCTGCTAGGTTGCAACCGATAGGTTGTAGAGTGTG 2580
Db 2315 ACTGATTCGACATTTCCAAAGCCTGCTGCTAGGTTGCAACCGATAGGTTGTAGAGTGTG 2374
Qy 2581 CAATACACTTGGTACAAATTTCAACCCCTTGGCAACTGCAAGCTTGGTGTGAAACAGCAT 2640
Db 2375 CAATACACTTGGTACAAATTTCAACCCCTTGGCAACTGCAAGCTTGGTGTGAAACAGCAT 2434
Qy 2641 CTTCAATTTCTGGAGCTCTTGTGCTGTATATCGACAGCCCAAGATCACTTGGGAAT 2700
Db 2435 CTTCAATTTCTGGAGCTCTTGTGCTGTATATCGACAGCCCAAGATCACTTGGGAAT 2494
Qy 2701 CAATACCAATGTTTCAGCTTGAGACAGAGGTCTGAGGCAACGAAATCTGGATCAGCGTATT 2760
Db 2495 CAATACCAATGTTTCAGCTTGAGACAGAGGTCTGAGGCAACGAAATCTGGATCAGCGTATT 2554
Qy 2761 TATCAGCAATTAATAGAACTTCAAGAGCCGAGAGGATGTCATATCTACACAGGCGTG 2820
Db 2555 TATCAGCAATTAATAGAACTTCAAGAGCCGAGAGGATGTCATATCTACACAGGCGTG 2614
Qy 2821 ATGTGTCATTTGAACCATCATCTTGGCAGCAGTAAACGAACTGCTTCTGGACCAATA 2880
Db 2615 ATGTGTCATTTGAACCATCATCTTGGCAGCAGTAAACGAACTGCTTCTGGACCAATA 2674
Qy 2881 TTTTGTCACTTAGGAACAGTTTCTGTTCCGTAAGCCATAGACGCTACTCTCGGCGC 2940

Db 2675 TTTTGTCACTTAGGAACAGTTTCTGTTCCGTAAGCCATAGCAGCTACTGCTCGGCGC 2734
Qy 2941 CTCTGTAGCACATACACTTAGCACCAACCTTGTGGGCAACGATAGATGACTTCTGGGG 3000
Db 2735 CTCTGTAGCACATACACTTAGCACCAACCTTGTGGGCAACGATAGATGACTTCTGGGG 2794
Qy 3001 TAAGGTTACCATCTCTTCTTAGTGGAGATGCAAAAACAAATTTCTTGGCAACGAACTT 3060
Db 2795 TAAGGTTACCATCTCTCTTAGTGGAGATGCAAAAACAAATTTCTTGGCAACGAACTT 2854
Qy 3061 TGGCAGGAACACCCAGCATCAGGGAAGTGGAGGCAAGATTTGCGTTCCACACGGAATAT 3120
Db 2855 TGGCAGGAACACCCAGCATCAGGGAAGTGGAGGCAAGATTTGCGTTCCACACGGAATAT 2914
Qy 3121 AGAGGCCAATCTTCTCAATAGTCTTGGAAAACGAGAGAGACTACACAGGGCAAGTCT 3180
Db 2915 AGAGGCCAATCTTCTCAATAGTCTTGGAAAACGAGAGAGACTACACAGGGCAAGTCT 2874
Qy 3181 CAACTTGAACAGTCTCGTCTAGTGTAGTTCATGGAATTTCTGAGCTTATCTATAGAGA 3240
Db 2975 CAACTTGAACAGTCTCGTCTAGTGTAGTTCATGGAATTTCTGAGCTTATCTATAGAGA 3034
Qy 3241 GATCAATGGCTCTCTTAACTGATTAATCTGGCAATTTGCAATTAAGTCTCTCTGGAAAGGAGCTT 3300
Db 3035 GATCAATGGCTCTCTTAACTGATTAATCTGGCAATTTGCAATTAAGTCTCTCTGGAAAGGAGCTT 3094
Qy 3301 CTAACACAGTGTCTTCAAGGAGCTCCATCAAACTTGGCAGTTAGTCTTAAAGGGCTT 3360
Db 3095 CTAACACAGTGTCTTCAAGGAGCTCCATCAAACTTGGCAGTTAGTCTTAAAGGGCTT 3154
Qy 3361 TGTCCACATTTTGAAGCAATTTGTGCAATTTGTTTGAATTTGCAATTTCCATTAATCTGTTCOG 3420
Db 3155 TGTCCACATTTTGAAGCAATTTGTGCAATTTGTTTGAATTTGCAATTTCCATTAATCTGTTCOG 3214
Qy 3421 TTTTCTGGATAGACAGAGAGGCACTTCAATTTCTTGTGAGGAGGCTTAGAAAGCT 3480
Db 3215 TTTTCTGGATAGACAGAGAGGCACTTCAATTTCTTGTGAGGAGGCTTAGAAAGCT 3274
Qy 3481 CAATTTGCACAATTTCAATACGACCTTCAAGAGGAGCTTCTTGTAGTTTGGATCTTCTT 3540
Db 3275 CAATTTGCACAATTTCAATACGACCTTCAAGAGGAGCTTCTTGTAGTTTGGATCTTCTT 3334
Qy 3541 TAGTTTGTCTTGTGTATCTCTGGCTTGGCAATCTCTTCTCTTCTTGTAGTGAACCTTTAGGG 3600
Db 3335 TAGTTTGTCTTGTGTATCTCTGGCTTGGCAATCTCTTCTCTTCTTGTAGTGAACCTTTAGGG 3394
Qy 3601 ACTTCATATCCAGGTTTCTCTCACCTCTGCAACGTCACCGTACTTGGCAGCTTAA 3660
Db 3395 ACTTCATATCCAGGTTTCTCTCACCTCTGCAACGTCACCGTACTTGGCAGCTTAA 3454
Qy 3661 CTAATGCAAAATAAAATAAGTCAGCAATTTCCAGGCTATATCTTCTTGGATTTAGCTT 3720
Db 3455 CTAATGCAAAATAAAATAAGTCAGCAATTTCCAGGCTATATCTTCTTGGATTTAGCTT 3514
Qy 3721 CTGCAAGTTTCAGCTTCTCTCCCTAATTTTAGCGTTCAACAAACCTTCTGTCGTAATA 3780
Db 3515 CTGCAAGTTTCAGCTTCTCTCCCTAATTTTAGCGTTCAACAAACCTTCTGTCGTAATA 3574
Qy 3781 ACCGTTTGTGTATAAAGAACCTTCTGGAGCATTTCTTACGATCCCAAGGTTGGCTTCCA 3840
Db 3575 ACCGTTTGTGTATAAAGAACCTTCTGGAGCATTTCTTACGATCCCAAGGTTGGCTTCCA 3634
Qy 3841 TGGCTCTAAGACCTTTTGAATGGCCAAAACAGGAAGTGGCTTCCAAAGTGAAGAACCAA 3900
Db 3635 TGGCTCTAAGACCTTTTGAATGGCCAAAACAGGAAGTGGCTTCCAAAGTGAAGAACCAA 3694
Qy 3901 CACCTGTTTGTCAACACCAAAATTTTCAAGCAGTCTCCATCAAACTCCAAATTCGATACCCA 3960
Db 3695 CACCTGTTTGTCAACACCAAAATTTTCAAGCAGTCTCCATCAAACTCCAAATTCGATACCCA 3754
Qy 3961 GCACTTTTGTGTCCTCAGATGTAGCACCTTTTATACCAAAACCGTGACGACGAGATT 4020
Db 3755 GCACTTTTGTGTCCTCAGATGTAGCACCTTTTATACCAAAACCGTGACGACGAGATT 3814

QY 4021 GGTAGACTCCAGTTTGTCTTATAGCCTCCGGAATAGACTTTTGGACGAGTACACCA 4080
DB |||||
DB 3815 GGTAGACTCCAGTTTGTCTTATAGCCTCCGGAATAGACTTTTGGACGAGTACACCA 3874
QY 4081 GGGCCACGAGTAATTTAGAGAGTACGACCAAAAGTAGTGAATAGACCATCGGGGGGT 4140
DB |||||
DB 3875 GGGCCACGAGTAATTTAGAGAGTACGACCAAAAGTAGTGAATAGACCATCGGGGGGT 3934
QY 4141 CAGTAGTCAAGACGCGCAAAATTTCACTGACAGGGAATTTTGGACATCTTTCAGAAA 4200
DB |||||
DB 3935 CAGTAGTCAAGACGCGCAAAATTTCACTGACAGGGAATTTTGGACATCTTTCAGAAA 3994
QY 4201 GTTCGTATTTCAGTAGTCAATTTCCGAGGNTCAATTAATGGGATATATACAGAGCAACAG 4260
DB |||||
DB 3995 GTTCGTATTTCAGTAGTCAATTTCCGAGGNTCAATTAATGGGATATATACAGAGCAACAG 4054
QY 4261 TGGAGTCAATCTACCAACTTTTGGGTTCTCAGAAAAAGCATAAACAGTTCTACTACCGC 4320
DB |||||
DB 4055 TGGAGTCAATCTACCAACTTTTGGGTTCTCAGAAAAAGCATAAACAGTTCTACTACCGC 4114
QY 4321 CATTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAAAGGCACAGCGATCTAGCAT 4380
DB |||||
DB 4115 CATTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAAAGGCACAGCGATCTAGCAT 4174
QY 4381 TAGCGGGCAAGGATGCAACTTTTATCAACAGGGTCTTATAGATAAACCCTAGCGCTGGGA 4440
DB |||||
DB 4175 TAGCGGGCAAGGATGCAACTTTTATCAACAGGGTCTTATAGATAAACCCTAGCGCTGGGA 4234
QY 4441 TATTCCTTTGGACAACTCTTTCGCAAAATCTAGGTCCAAATCACTTCAATGATACCAT 4500
DB |||||
DB 4235 TATTCCTTTGGACAACTCTTTCGCAAAATCTAGGTCCAAATCACTTCAATGATACCAT 4294
QY 4501 TATTCCTTTGGACAACTCTTTCGCAAAATCTAGGTCCAAATCACTTCAATGATACCAT 4560
DB |||||
DB 4295 TATTCCTTTGGACAACTCTTTCGCAAAATCTAGGTCCAAATCACTTCAATGATACCAT 4354
QY 4561 ACTCAACTTGCATTAATTTGAAGCTCAGTGCATTTGAGTGAACCTTGAATGAGTGTGCA 4620
DB |||||
DB 4355 ACTCAACTTGCATTAATTTGAAGCTCAGTGCATTTGAGTGAACCTTGAATGAGTGTGCA 4414
QY 4621 GCTGTGACAGCAGTAGGGAACACGGCTTTTCTACCAAACTCAAGAAATTTATCAAACT 4680
DB |||||
DB 4415 GCTGTGACAGCAGTAGGGAACACGGCTTTTCTACCAAACTCAAGAAATTTATCAAACT 4474
QY 4681 CTGCAACACTTCCGTATGCGAGTAGCAAGGGAATGTCATCTACTTGAAGTGGACAGTGAG 4740
DB |||||
DB 4475 CTGCAACACTTCCGTATGCGAGTAGCAAGGGAATGTCATCTACTTGAAGTGGACAGTGAG 4534
QY 4741 TGTAGTCTTGAAGAAATCTGAAGCCGTATTTTATATCACTGAGTCACTCAGTCAAGAGA 4800
DB |||||
DB 4535 TGTAGTCTTGAAGAAATCTGAAGCCGTATTTTATATCACTGAGTCACTCAGTCAAGAGA 4594
QY 4801 TCCTCTACCCCGACGCAATCGTGGCCGA 4828
DB |||||
DB 4595 TCCTCTACCCCGACGCAATCGTGGCCGA 4654
QY 4829 ----- 4828
DB 4655 CTCGTGAGAGAGGTGTTGCTGACTCATACAGGCTGNAATCGCCCCATCATCCAGCCAGA 4714
QY 4829 ----- 4828
DB 4715 AAGTGAGGAGCCACGGTTGATGAGAGCTTTGTTGTAGGTGACCACTGTTGTTGATTTGA 4774
QY 4829 ----- 4828
DB 4775 ACTTTTGTCTTGGCACGGAACGGTCTGGTGTGTCGGGAGAGATCGGTGATCTGATCTCTCA 4834
QY 4829 ----- 4828
DB 4835 ACTCAGAAAAAGTTCGATTTTATTTCAACAAAGCCCGCTCCCGTCAAGTCAAGCGTAATGCT 4894

QY 4829 ----- 4828
DB 4895 CTGCCAGTGTACAAACCAATTTAACCAATTTCTGATTAGAAAACTCATCGAGCATCAAAATG 4954
QY 4829 ----- 4828
DB 4955 AAACCTGCAATTTTATTCATATCAGGATTTCAATPACCATATTTTGGAAAAAGCGCTTTCTG 5014
QY 4829 ----- 4828
DB 5015 TAATGAGGAGAAAACTCACGAGGCGAGTTCCATAGATGGCAAGATCCTGTGTATCGGTC 5074
QY 4829 ----- 4828
DB 5075 TCGGATCCGACTCGTCCACATCAATCAACCTTTAAATTTCCCTCGTCAAAAAATAAG 5134
QY 4829 ----- 4828
DB 5135 GTTATCAAGTGAAGAAATCACCATGAGTGACGACTGATCCGGTGAGAAATGGCAAAAGCTT 5194
QY 4829 ----- 4828
DB 5195 ATGCATTTCTTTCCAGACTTGTTCACAGGCCAGCCATTACGCTCGTCAATAAATCACT 5254
QY 4829 ----- 4828
DB 5255 CGCATCAACCAACCGTTTATTTCTGTTGATTGGCTGAGCGAGAGAAATACGCGATC 5314
QY 4829 ----- 4828
DB 5315 GCTGTAAAGGACAAATTAACAACAGGAATCGAATGGAACCGCGGAGGAAACATCCGAG 5374
QY 4829 ----- 4828
DB 5375 CGCATCAACCAATTTTCTCACTGAAATCAGGATATTTCTTCTAATACCTGGAATGCTGTTT 5434
QY 4829 ----- 4828
DB 5435 CCGGGGATCGCAGTGTGATGAACCAATGATCATCAGGAGTACGGAATAAAATGCTTGAT 5494
QY 4829 ----- 4828
DB 5495 GGTGGAAGAGGCATAAATCCCGTCAGCCAGTTTGTCTGACCATCTCATCTGTAACATC 5554
QY 4829 ----- 4828
DB 5555 ATTGGCAACGCTACCTTTGCAATGTTTCAGAAAAAACTCTGCGGCATCGGGCTTTCCCAT 5614
QY 4829 ----- 4828
DB 5615 CAATCGATAGATTGTGCACTCGACCTGATTTGCCGACATTTATCGGAGCCCATTTATACCAT 5674
QY 4829 ----- 4828
DB 5675 TAAATCAGATCCATGTTGGAATTTAATCGCGGCTCGAGCAAGAGCTTTCCGTTGAAT 5734
QY 4829 ----- 4828
DB 5735 ATGGCTATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTATTGTTTCATGA 5794
QY 4829 ----- 4828
DB 5795 TGATATATTTTATCTGTGCAATGTAACATCAGAGATTTGAGACACAACCTGGCTTC 5854
QY 4829 ----- 4828
DB 5855 CCCCCCCCCCTGCGAGTGTGCAATCCCGGCGCATCCCGGCGCATCAGGTTGCTGCGGCTATAT 5914
QY 4880 CGCGCATCAGGATGGGGAAGATCGGCTCGCACTTTCGGGCTCATGAGCGCTGTTT 4939
DB |||||
DB 5915 CGCGCATCAGGATGGGGAAGATCGGCTCGCACTTTCGGGCTCATGAGCGCTGTTT 5974
QY 4940 CGCGTGGGTATGGTGGCAGGCCCTCGTGGCGGGGACTGTTGGGCGGCATCTCCTTGCA 4999


```
QY 7160 CTACGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGCGAGACCCAC 7219
Db 8195 CTACGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGCGAGACCCAC 8254
QY 7220 GCTCACGGCTCCAGATTTATCAGCAATAAACAGCCAGCGGAGCGGCGAGCGAGAA 7279
Db 8255 GCTCACGGCTCCAGATTTATCAGCAATAAACAGCCAGCGGAGCGGCGAGCGAGAA 8314
QY 7280 GTGCTCCTGCAACTTTATCCGCTCCATCCAGTCTATTAAATTTTCCGGGAAGCTAGAG 7339
Db 8315 GTGCTCCTGCAACTTTATCCGCTCCATCCAGTCTATTAAATTTTCCGGGAAGCTAGAG 8374
QY 7340 TAAGTAGTCCCGAGTTAATAGTTTGGCAACGTTGTGTCATTTGTCGAGGCGATCGTG 7399
Db 8375 TAAGTAGTCCCGAGTTAATAGTTTGGCAACGTTGTGTCATTTGTCGAGGCGATCGTG 8434
QY 7400 TGTCAAGTAAAGTTGGTATGGCTTCATTCAGCTCCGTTCCCAACGATCAAGCGAG 7459
Db 8435 TGTCAAGTAAAGTTGGTATGGCTTCATTCAGCTCCGTTCCCAACGATCAAGCGAG 8494
QY 7460 TTAATGATCCCGCATGTTGTGCAAAAAGGGTTAGTCTCTTGGTCTCCGATCGTTG 7519
Db 8495 TTAATGATCCCGCATGTTGTGCAAAAAGGGTTAGTCTCTTGGTCTCCGATCGTTG 8554
QY 7520 TCAGAAGTAAAGTTGGTATGGCTTCATTCAGCTCCGTTCCCAACGATCAAGCGAG 7579
Db 8555 TCAGAAGTAAAGTTGGTATGGCTTCATTCAGCTCCGTTCCCAACGATCAAGCGAG 8614
QY 7580 TTAATGATCCCGCATGTTGTGCAAAAAGGGTTAGTCTCTTGGTCTCCGATCGTTG 7639
Db 8615 TTAATGATCCCGCATGTTGTGCAAAAAGGGTTAGTCTCTTGGTCTCCGATCGTTG 8674
QY 7640 TCTCAGATAGTGTATCGGCGACCGAGTTGCTCTCCCGCGGTCAACCGGATTAATA 7699
Db 8675 TCTCAGATAGTGTATCGGCGACCGAGTTGCTCTCCCGCGGTCAACCGGATTAATA 8734
QY 7700 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAACAGTTTCTCGGGCGAA 7759
Db 8735 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAACAGTTTCTCGGGCGAA 8794
QY 7760 AACTCTCAAGATCTTACCGCTGTGTAGATCCAGTTCAGTGTACCCACTCGTGCACCCA 7819
Db 8795 AACTCTCAAGATCTTACCGCTGTGTAGATCCAGTTCAGTGTACCCACTCGTGCACCCA 8854
QY 7820 ACTGATCTTACGATCTTTTACTTTTCAACAGCGTTTCTGGGTGAGCAAAACAGGAAGGC 7879
Db 8895 ACTGATCTTACGATCTTTTACTTTTCAACAGCGTTTCTGGGTGAGCAAAACAGGAAGGC 8914
QY 7880 AAAATGCGGCAAAAAGGAATAAGGCGGACACGGAATGTTGAATCTCATACTCTTCC 7939
Db 8915 AAAATGCGGCAAAAAGGAATAAGGCGGACACGGAATGTTGAATCTCATACTCTTCC 8974
QY 7940 TTTTTCATATTTTGAAGCATTTTATCAGGGTTATTTCTCTCATGCGGATACATATTTG 7999
Db 8975 TTTTTCATATTTTGAAGCATTTTATCAGGGTTATTTCTCTCATGCGGATACATATTTG 9034
QY 8000 AATGTATTTAGAAAAATAAACAATAGGGTTCCGCGCAATTTCCCGGAAAAGTGGCCAC 8059
Db 9035 AATGTATTTAGAAAAATAAACAATAGGGTTCCGCGCAATTTCCCGGAAAAGTGGCCAC 9094
QY 8060 CTGACGCTTAGAAACCATTTATTCATGACATTAACCTATAAATAAGGCGTATCACGA 8119
Db 9095 CTGACGCTTAGAAACCATTTATTCATGACATTAACCTATAAATAAGGCGTATCACGA 9154
QY 8120 GGCCCTTTTCTGCTTCAAGAATAAATCTCATGTTTGAAGCTTATCATCATGAAGCTGAC 8179
Db 9155 GGCCCTTTTCTGCTTCAAGAATAAATCTCATGTTTGAAGCTTATCATCATGAAGCTGAC 9214
QY 8180 TCATGTTGGTATTTGAAATAGACGCGAGTCGGGAACACTGGAATAAAGTATTAAT 8239
Db 9215 TCATGTTGGTATTTGAAATAGACGCGAGTCGGGAACACTGGAATAAAGTATTAAT 9274
```

```
QY 8240 CG 8241
Db 9275 CG 9276

RESULT 6
US-10-104-271-11
; Sequence 11, Application US/10104271
; Publication No. US20030181658A1
; GENERAL INFORMATION:
; APPLICANT: Madison, Edwin
; APPLICANT: Jiunn-Chern Yeh
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING SERINE PROTEASE CVSP14, THE ENCODED
; FILE REFERENCE: 24745-1614
; TITLE OF INVENTION: POLYPEPTIDES AND METHODS BASED THEREON
; CURRENT APPLICATION NUMBER: US/10/104,271
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/278,166
; PRIOR FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 9276
; TYPE: DNA
; ORGANISM: Pichia pastoris
US-10-104-271-11

Query Match 79.6%; Score 6561.6; DB 15; Length 9276;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;

QY 1 AGATCTAACTCCAAAGACGAAAGGTTGAATGAAACCTTTTGGCCATCCGACATCCACAG 60
Db 1 AGATCTAACTCCAAAGACGAAAGGTTGAATGAAACCTTTTGGCCATCCGACATCCACAG 60
QY 61 GTCCATTTCTCACATAGTCCCAAGACGACAGAGGGATACACTAGCAGCAGACCGT 120
Db 61 GTCCATTTCTCACATAGTCCCAAGACGACAGAGGGATACACTAGCAGCAGACCGT 120
QY 121 TGCAACAGCAGGACCTCCACCTCTTCTCTCCTCAACACCCACTTTTGCATTCGAAAAACC 180
Db 121 TGCAACAGCAGGACCTCCACCTCTTCTCTCCTCAACACCCACTTTTGCATTCGAAAAACC 180
QY 181 AGCCAGTATTTGGCTTGTATGAGTCCGCTCATTCCTCAATTCCTTCTATTAGGCTACTA 240
Db 181 AGCCAGTATTTGGCTTGTATGAGTCCGCTCATTCCTCAATTCCTTCTATTAGGCTACTA 240
QY 241 ACACCATGACTTTTATTAGCTGTCTATCTGGCCCCCTCTGGCGAGGTTTCATGTTTGTTA 300
Db 241 ACACCATGACTTTTATTAGCTGTCTATCTGGCCCCCTCTGGCGAGGTTTCATGTTTGTTA 300
QY 301 TTTCGAAATGCAACAGCTCCGATTTACACCCGAAACATCACTCCAGATGAGGCTTTCTG 360
Db 301 TTTCGAAATGCAACAGCTCCGATTTACACCCGAAACATCACTCCAGATGAGGCTTTCTG 360
QY 361 AGTGTGGGTCAAAATAGTTTTCATGTTCCCAAAATGGCCCAAACTGCACAGTTTAAACGCT 420
Db 361 AGTGTGGGTCAAAATAGTTTTCATGTTCCCAAAATGGCCCAAACTGCACAGTTTAAACGCT 420
QY 421 GTCTTGGAACTTAATATGACAAAAAGCGTGTCTCATCCAGATGAACCTAGTTTGGTTCG 480
Db 421 GTCTTGGAACTTAATATGACAAAAAGCGTGTCTCATCCAGATGAACCTAGTTTGGTTCG 480
QY 481 TTGAATGCTTAACGCGCAGTTGGTCAAAAAGAACTTCCAAAAGTCCGATACCGTTGT 540
Db 481 TTGAATGCTTAACGCGCAGTTGGTCAAAAAGAACTTCCAAAAGTCCGATACCGTTGT 540
QY 541 CTGTGTTGGTATTTGATGACGAATGCTCAAAAAATAATCTCAATTAATGCTTAGCCAGTCT 600
Db 541 CTGTGTTGGTATTTGATGACGAATGCTCAAAAAATAATCTCAATTAATGCTTAGCCAGTCT 600
QY 601 CTCTATCGCTTCTGAAACCGCGTGCATCTGTGCGAAACGCAATGGGAAACACCGCT 660
Db 601 CTCTATCGCTTCTGAAACCGCGTGCATCTGTGCGAAACGCAATGGGAAACACCGCT 660
```


QY 2821 ATGTGTCATTTTGAACCATCATCTTGGCAGCAGTAACGAATCGTTTCTTGGAACCAATA 2880
Db 2615 ATGTGTCATTTTGAACCATCATCTTGGCAGCAGTAACGAATCGTTTCTTGGAACCAATA 2674
QY 2881 TTTTGTACACACTTAGGAACAGTTTCTGTTCCGTAAGCCATPAGCAGCTACTGCTCGGCGC 2940
Db 2675 TTTTGTACACACTTAGGAACAGTTTCTGTTCCGTAAGCCATPAGCAGCTACTGCTCGGCGC 2734
QY 2841 CTCCTGTAGCAGATACACTTAGCACCACCTTGTGGCAAGTAGAGTACTTCTGGGG 3000
Db 2735 CTCCTGTAGCAGATACACTTAGCACCACCTTGTGGCAAGTAGAGTACTTCTGGGG 2794
QY 3001 TAAGGTTACCATCTCTTCTAGGTGGAGATGCAAAAACAAATTTCTTTCGCAACCAAGCAACTT 3060
Db 2795 TAAGGTTACCATCTCTTCTAGGTGGAGATGCAAAAACAAATTTCTTTCGCAACCAAGCAACTT 2854
QY 3061 TGGCAGGAACACCCAGCATACGGGAAGTGGAAAGCAGAAATGGGGTCCACCAAGGATAT 3120
Db 2855 TGGCAGGAACACCCAGCATACGGGAAGTGGAAAGCAGAAATGGGGTCCACCAAGGATAT 2914
QY 3121 AGAGGCCAACTTCTCAATAGTCTTGCAAAACGAGAGCAGACTACACCAAGGCAAGTCT 3180
Db 2915 AGAGGCCAACTTCTCAATAGTCTTGCAAAACGAGAGCAGACTACACCAAGGCAAGTCT 2974
QY 3181 CAACCTTGCAACGCTCTCCGTTAGTTAGCTTCATGGAATTTCTGAGCTTATCTATAGAGA 3240
Db 2975 CAACCTTGCAACGCTCTCCGTTAGTTAGCTTCATGGAATTTCTGAGCTTATCTATAGAGA 3034
QY 3241 GATCAATGGCTCTCTTAAGCTTATCTGCAATTTGCAATTAAGTCTCTCTGGGAAAGGAGCTT 3300
Db 3035 GATCAATGGCTCTCTTAAGCTTATCTGCAATTTGCAATTAAGTCTCTCTGGGAAAGGAGCTT 3094
QY 3301 CTAACACAGGCTGTCTCAAAGCGACTCCATCAAACTTGGCAGTAGTTCTPAAAAGGGCTT 3360
Db 3095 CTAACACAGGCTGTCTCAAAGCGACTCCATCAAACTTGGCAGTAGTTCTPAAAAGGGCTT 3154
QY 3361 TGTCAACCTTTTGAAGCAATTTGCGCAATTTGGTTGACTTAATTCATATCTGTTCCG 3420
Db 3155 TGTCAACCTTTTGAAGCAATTTGCGCAATTTGGTTGACTTAATTCATATCTGTTCCG 3214
QY 3421 TTTTCTGATAGGACGACGAAGGCACTTCAATTTCTTGTGAGGAGGCTTGAAGAACGT 3480
Db 3215 TTTTCTGATAGGACGACGAAGGCACTTCAATTTCTTGTGAGGAGGCTTGAAGAACGT 3274
QY 3481 CAATTTTGCACAATTCATAGCACTTCAGAGGCACTTCTTTAGTTGGATCTCTCTT 3540
Db 3275 CAATTTTGCACAATTCATAGCACTTCAGAGGCACTTCTTTAGTTGGATCTCTCTT 3334
QY 3541 TAGGTTGTTCTGTTGATCTCTGCTTGGCATCTCTCTTCTTCTAGTGACCTTTAGGG 3600
Db 3335 TAGGTTGTTCTGTTGATCTCTGCTTGGCATCTCTCTTCTTCTAGTGACCTTTAGGG 3394
QY 3601 ACTTCATATCCAGGTTCTCTCACTGCTGTCACCAAGTCAACCGTACTTGGCACTATA 3660
Db 3395 ACTTCATATCCAGGTTCTCTCACTGCTGTCACCAAGTCAACCGTACTTGGCACTATA 3454
QY 3661 CTAATGCAAAATAAATAAGTCACCACTTCCAGGCTATATCTCTCTGATTTAGCTT 3720
Db 3455 CTAATGCAAAATAAATAAGTCACCACTTCCAGGCTATATCTCTCTGATTTAGCTT 3514
QY 3721 CTGCAAGTTTCATGAGTTCTCTCCCTAAATTTTGTAGGTTCAACAAACTTCGTCGCAATA 3780
Db 3515 CTGCAAGTTTCATGAGTTCTCTCCCTAAATTTTGTAGGTTCAACAAACTTCGTCGCAATA 3574
QY 3781 ACCGTTTGGTATAGAACCTTCTGGAGCATTTGCTTTAGCATCCCAAGGTGGCTTCCA 3840
Db 3575 ACCGTTTGGTATAGAACCTTCTGGAGCATTTGCTTTAGCATCCCAAGGTGGCTTCCA 3634
QY 3841 TGGCTCTAAGACCTTTGATTTGGCCAAAACAGGAAGTGGCTTCCAAAGTGAAGAAACCAA 3900
Db 3635 TGGCTCTAAGACCTTTGATTTGGCCAAAACAGGAAGTGGCTTCCAAAGTGAAGAAACCAA 3694
QY 3901 CACCTGTTTGTTCACCAACCAAAATTTCAAGCAGTCTCCATCACATCCAAATTCGNATCCCA 3960

Db 3695 CACCTGTTTGTTCACCAACCAAAATTTCAAGCAGTCTCCATCACAAATCCAAATTCGNATCCCA 3754
QY 3961 GCAACTTTTGAAGTTGCTCCAGATCTAGCAGCTTATACCAAAACCTGTAGCAGAGATT 4020
Db 3755 GCAACTTTTGAAGTTGCTCCAGATCTAGCAGCTTATACCAAAACCTGTAGCAGAGATT 3814
QY 4021 GGTAGACTCCCAAGTTTGTCTTATAGCTCTCCGGAATAGACTTTTTCGACAGGATACCA 4080
Db 3815 GGTAGACTCCCAAGTTTGTCTTATAGCTCTCCGGAATAGACTTTTTCGACAGGATACCA 3874
QY 4081 GGCCCAACGAGTAATTAGAGAGTCAAGCCCAAGTAGTCAATAGACCACTCCGGGGCGT 4140
Db 3875 GGCCCAACGAGTAATTAGAGAGTCAAGCCCAAGTAGTCAATAGACCACTCCGGGGCGT 3934
QY 4141 CAGTAGTCAAAAGACGCAACAAATTTTCACTGACAGGGAATTTTTCACATCTTTCAGAAA 4200
Db 3935 CAGTAGTCAAAAGACGCAACAAATTTTCACTGACAGGGAATTTTTCACATCTTTCAGAAA 3994
QY 4201 GTTCGTATTCAGTAGTCAATTTGCCGAGCATCAATTAATGGGGATTTATACCAAGACCAAG 4260
Db 3995 GTTCGTATTCAGTAGTCAATTTGCCGAGCATCAATTAATGGGGATTTATACCAAGACCAAG 4054
QY 4261 TGGAGTCAACATCTACCAACTTTGCGGCTTCAGAAAAGCATAAACAGTTCTACTACCGC 4320
Db 4055 TGGAGTCAACATCTACCAACTTTGCGGCTTCAGAAAAGCATAAACAGTTCTACTACCGC 4114
QY 4321 CATTAGTGAATTTTCAATTCGCCAGTGAGAGAAAAGGCAAGCATCTACTACCGC 4380
Db 4115 CATTAGTGAATTTTCAATTCGCCAGTGAGAGAAAAGGCAAGCATCTACTACCGC 4174
QY 4381 TAGCGGCAAGGATGCAACTTTATCAACCGGCTCTATAGATAACCTAGCGCTCGGA 4440
Db 4175 TAGCGGCAAGGATGCAACTTTATCAACCGGCTCTATAGATAACCTAGCGCTCGGA 4234
QY 4441 TCATCTCTTTGACAACTCTTTCTGCCAAATCTAGTCCAAAATCACTTCATTTGATACCAT 4500
Db 4235 TCATCTCTTTGACAACTCTTTCTGCCAAATCTAGTCCAAAATCACTTCATTTGATACCAT 4294
QY 4501 TATTGTCAACTTTGAGGAGTGTGCTAGCTCTCAAAATTTGGTCTCTGTACGGATG 4560
Db 4295 TATTGTCAACTTTGAGGAGTGTGCTAGCTCTCAAAATTTGGTCTCTGTACGGATG 4354
QY 4561 ACTCAACTTTGACAACTTTAAAGTCTAGTCAAGTTGAGTCAAACTTTGATCAGGTTGTGCA 4620
Db 4355 ACTCAACTTTGACAACTTTAAAGTCTAGTCAAGTTGAGTCAAACTTTGATCAGGTTGTGCA 4414
QY 4621 GCTGTGACGAGCATPAGGAAACACGGCTTTTCTACCAAACTCAAGGAATTTCAAACT 4680
Db 4415 GCTGTGACGAGCATPAGGAAACACGGCTTTTCTACCAAACTCAAGGAATTTCAAACT 4474
QY 4681 CTGCAAACTTTGCGTATGAGTCAAGGGAATTTGATCACTTTGAAAGTCCGACAGTGAG 4740
Db 4475 CTGCAAACTTTGCGTATGAGTCAAGGGAATTTGATCACTTTGAAAGTCCGACAGTGAG 4534
QY 4741 TGTAGTCTTGAGAAATTTCTGAAGCGGATTTTATTTATCAGTGAAGTCAATCAGGAGA 4800
Db 4535 TGTAGTCTTGAGAAATTTCTGAAGCGGATTTTATTTATCAGTGAAGTCAATCAGGAGA 4594
QY 4801 TCCTCTACCGCGGAGCATCTGGCCGACCTGACAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 4828
Db 4595 TCCTCTACCGCGGAGCATCTGGCCGACCTGACAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 4654
QY 4829 TCCTCTACCGCGGAGCATCTGGCCGACCTGACAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 4828
Db 4655 CTCGTGAAGAGGTTGTTCTGACTCATACAGGCGTGAATCGCCCATCATCCAGCCAGA 4714
QY 4829 TCCTCTACCGCGGAGCATCTGGCCGACCTGACAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 4828
Db 4715 AAGTGAGGAGCCAGGTTGATGAGAGCTTTGTTGAGGTGGACCAAGTTGGTGAATTTGA 4774
QY 4829 TCCTCTACCGCGGAGCATCTGGCCGACCTGACAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 4828

Db	4775	ACTTTTGCTTGGCCACGGAACGGTCTCGGTTGTCGGGAAGATCGGTGATCTGATCCCTTCA	4834
Qy	4829	-----	4828
Db	4835	ACTCAGCAAAAGTTTCGATTTATTTCAACAAAGCCGCTCCGTCAGTCAAGCTAATGCT	4894
Qy	4829	-----	4828
Db	4895	CTGCCAGTGTTPACAAACCAATTAACCAATTCGTATTAGAAAACTCATCGAGCATCAAAATG	4954
Qy	4829	-----	4828
Db	4955	AAACTGCAATTTATTCATATCAGGATTAACAATACCATAATTTTGA AAAAGCCGTTTCTG	5014
Qy	4829	-----	4828
Db	5015	TAATGAAGGAAAACTCACGAGGAGTTTCCATAGGATGCAAGATTCCTGGTATCGGTC	5074
Qy	4829	-----	4828
Db	5075	TGCGATTCGGAATCGTCCAAATCAATACAACCTATTAAATTTCCCTCGTCMAAAATAAG	5134
Qy	4829	-----	4828
Db	5135	GTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGCAAAAAGCTT	5194
Qy	4829	-----	4828
Db	5195	ATGCATTTCTTCCAGACTTGTTCACACAGGCCCAATAGGCTCGTCATCAAAATCACT	5254
Qy	4829	-----	4828
Db	5255	CGCATCAACCAACGGTTATTTCATTCGTGATTGCGCTGAGCGAGACGAAATACGGATC	5314
Qy	4829	-----	4828
Db	5315	GCTGTTAAAGGACAATTAACAAACAGGAATCGAATGCAACGGCGCAGGAACAACCTGCCAG	5374
Qy	4829	-----	4828
Db	5375	CGCATCAACAATATTTTCACTCGAATCAGGATATCTTCTAATACCTGGAATGCTGTTT	5434
Qy	4829	-----	4828
Db	5435	CCCGGGATCGAGTGTGAGTAACCATGCATCATCAGAGTAGCGGATATAAAATGCTTCAT	5494
Qy	4829	-----	4828
Db	5495	GGTGGAAAGGCATAAAATTCGGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAAACATC	5554
Qy	4829	-----	4828
Db	5555	ATTGCAACGCTACCTTTGCCATGTTTCAGAAACAACCTCGCGCATCGGGCTTCCCAT	5614
Qy	4829	-----	4828
Db	5615	CAATCATAGATGTGCAACCTGATGTCGCCAGACATTATCGAGGCCAATTTATACCCATA	5674
Qy	4829	-----	4828
Db	5675	TAAATCAGCATCCATGTTGGAATTTAATCGCGGCTCGAGCAAGACGTTTCCCGTTGAAT	5734
Qy	4829	-----	4828
Db	5735	ATGGCTCATACACCCCTTGTATTACTGTTATGTAAGCAGACAGTTTATTGTTTCATGA	5794
Qy	4829	-----	4828
Db	5795	TGATATATTTTATCTTGCAATGTAAATCAATCAGAGATTTTGAGACACAACGTGGCTTTC	5854
Qy	4829	-----	4828
Db	5855	CCCCCCCCCTCTGCAAGTTCGGCATCAACCGGCGCCACAGGTGCGGTTGCTGCGGCTATAT	5914

Qy	4880	CGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTT	4939
Db	5915	CGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTT	5974
Qy	4940	CGCGGTGGGTATGGTGGCAGGCCCCGTCGCCGGGGAAGTGTGGGCGGCATCTCCTTGCA	4999
Db	5975	CGCGGTGGGTATGGTGGCAGGCCCCGTCGCCGGGGAAGTGTGGGCGGCATCTCCTTGCA	6034
Qy	5000	TGCAACATTCCTTGGGCGGCGGTCTCAACGGCCCTCAACCTACTACTCTGGGCTGCTTCT	5059
Db	6035	TGCAACATTCCTTGGGCGGCGGTCTCAACGGCCCTCAACCTACTACTCTGGGCTGCTTCT	6094
Qy	5060	AATGACAGAGTCGATATAGGAGAGCGTCGAGTATCTATGATTGGAAGTATGCGGAATGGT	5119
Db	6095	AATGACAGAGTCGATATAGGAGAGCGTCGAGTATCTATGATTGGAAGTATGCGGAATGGT	5154
Qy	5120	GATACCCGCAATTCCTTCAAGTCTCTTATCAGATTCCTTATCAGATTCCTTATCAGATTC	5179
Db	6155	GATACCCGCAATTCCTTCAAGTCTCTTATCAGATTCCTTATCAGATTCCTTATCAGATTC	6214
Qy	5180	CGGAGGAGAGATTCATGTTAAATTTCTCAGCTTTTGGTCTCATCAGTAGACTCGAACTG	5239
Db	6215	CGGAGGAGAGATTCATGTTAAATTTCTCAGCTTTTGGTCTCATCAGTAGACTCGAACTG	6274
Qy	5240	TGAGACTATCTCGGTTATGACAGCAAAATGTCCTTTTGGAGACAGTAAATGAAAGTCCC	5299
Db	6275	TGAGACTATCTCGGTTATGACAGCAAAATGTCCTTTTGGAGACAGTAAATGAAAGTCCC	6334
Qy	5300	ACCAATAAAGAAATTCCTTTTATCAGGAACAACTTTCTGTTTCGAACTTTTCGGTCCC	5359
Db	6335	ACCAATAAAGAAATTCCTTTTATCAGGAACAACTTTCTGTTTCGAACTTTTCGGTCCC	6394
Qy	5360	TTGAACTATAAATGATAGAGTGGATATGTCGGTAGGAATGAGCGGCAATGCTTACC	5419
Db	6395	TTGAACTATAAATGATAGAGTGGATATGTCGGTAGGAATGAGCGGCAATGCTTACC	6454
Qy	5420	TTCTGACCTTCAAGAGGTATGAGGTTTGTAGATCTATGATGCTTATGAGTGAACAC	5479
Db	6455	TTCTGACCTTCAAGAGGTATGAGGTTTGTAGATCTATGATGCTTATGAGTGAACAC	6514
Qy	5480	GTGCTATTTCTTCAACCAATTCGGAATCCAGAGAAATCAAGTTGTTGTTCTACTAT	5539
Db	6515	GTGCTATTTCTTCAACCAATTCGGAATCCAGAGAAATCAAGTTGTTGTTCTACTAT	6574
Qy	5540	GATCCAAAGCAGTCCGCTTGAACCTGCAATAGTGTCTGCTGCTGCTGCTGCTGCTGCTG	5599
Db	6575	GATCCAAAGCAGTCCGCTTGAACCTGCAATAGTGTCTGCTGCTGCTGCTGCTGCTGCTG	6634
Qy	5600	TGATGAATAAATCTAGTCTTTGATCTAAATTAATCTTGACGAGCCGAGGATTAATACC	5659
Db	6635	TGATGAATAAATCTAGTCTTTGATCTAAATTAATCTTGACGAGCCGAGGATTAATACC	6694
Qy	5660	CAATCTAAACCTCTTTTAAACGTTTAAAGCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTG	5719
Db	6695	CAATCTAAACCTCTTTTAAACGTTTAAAGCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTG	6754
Qy	5720	AAATCAGTCTGATCTGATCTCATCAATCTTGAGGCGCACTATCTTGTGTTAGAGAAAT	5779
Db	6755	AAATCAGTCTGATCTGATCTCATCAATCTTGAGGCGCACTATCTTGTGTTAGAGAAAT	6814
Qy	5780	TTGCGGAGATCGGATATCGAGAAAAGGTACGCTGATTTTAAACGTGAAATTTATCTCAA	5839
Db	6815	TTGCGGAGATCGGATATCGAGAAAAGGTACGCTGATTTTAAACGTGAAATTTATCTCAA	6874
Qy	5840	GATCTCTGCTCGCGGTTTTCGGTGATGACGGTGAACCTCTGACACATGCAAGTCCCG	5899
Db	6875	GATCTCTGCTCGCGGTTTTCGGTGATGACGGTGAACCTCTGACACATGCAAGTCCCG	6934
Qy	5900	GAGACGTCACAGCTTGTCTGTAAGCGGATGCGGAGCAGACAAGCCGTCAGGCGCG	5959
Db	6935	GAGACGTCACAGCTTGTCTGTAAGCGGATGCGGAGCAGACAAGCCGTCAGGCGCG	6994


```

Db      9155  GGCCCTTTTCGCTTCAAGAAATTAATCTTCATGTTTGCAGCTTATCATCGATAGCTGAC  9214
Qy      8180  TCATGTTGGTATTGTGAAATAGACGCAGATCGGGAAACACTCAAAAAATAACAGTTATTATT  8239
Db      9215  TCATGTTGGTATTGTGAAATAGACGCAGATCGGGAAACACTCAAAAAATAACAGTTATTATT  9274
Qy      8240  CG 8241
Db      9275  CG 9276

RESULT 7
US-10-147-211A-13
; Sequence 13, Application US/10147211A
; Publication No. US20030235900A1
; GENERAL INFORMATION:
; APPLICANT: Madison, Edward
; APPLICANT: Yeh, Jiumn-Chern
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING A TRANSMEMBRANE SERINE P
; TITLE OF INVENTION: ENCODED POLYPEPTIDES AND METHODS BASED THEREON
; FILE REFERENCE: 24745-1616
; CURRENT APPLICATION NUMBER: US/10/147,211A
; CURRENT FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 60/291,001
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 9276
; TYPE: DNA
; ORGANISM: Pichia pastoris
US-10-147-211A-13

Query Match          79.6%; Score 6561.6; DB 16; Length 9276;
Best Local Similarity 84.7%; Pred No. 0;
Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;

```

Query Match	79.6%;	Score 6561.6;	DB 16;	Length 9276;
Best Local Similarity	84.7%;	Ref. No. 0;		
Matches 8031;	Conservative 0;	Mismatches 4;	Indels 1447;	Gaps 27;
Qy	1	AGNCTAACATCCAAAGACGAAAGGTTCAATGAACCTTTTGGCATCCGACATCCACAG	60	
Db	1	AGATCTAACATCCAAAGACGAAAGGTTGAATGAACCTTTTGGCATCCGACATCCACAG	60	
Qy	61	GTCCATTTCTCACATAAGTGCACAAACGCAACAGAGGGGATACACTAGCAGCAGCGT	120	
Db	61	GTCCATTTCTCACATNAGTGCCAAACGCAACAGAGGGGATACACTAGCAGCAGCGT	120	
Qy	121	TGCAAAACGACGAGACTCCACATCCCTCTTCTCTCAACACGCCACTTTTGGCATCGAAAAACC	180	
Db	121	TGCAAAACGACGAGACTCCACATCCCTCTTCTCTCAACACGCCACTTTTGGCATCGAAAAACC	180	
Qy	181	AGCCAGTATTGGGCTTGATTGGAGTCGCTCATTTCCAAATTCCTTCTATTAGGCTACTA	240	
Db	181	AGCCAGTATTGGGCTTGATTGGAGTCGCTCATTTCCAAATTCCTTCTATTAGGCTACTA	240	
Qy	241	ACACCATGACTTTATTAGCCTGCTCTATCTGTGCGCCCTCTGGCGAGGTTTCATGTTGTTTA	300	
Db	241	ACACCATGACTTTATTAGCCTGCTCTATCTGTGCGCCCTCTGGCGAGGTTTCATGTTGTTTA	300	
Qy	301	TTTCCGAATGCAACAGCTCCGATTTACCCGGAACATCACTCCAGATGAGGGCTTCTTG	360	
Db	301	TTTCCGAATGCAACAGCTCCGATTTACACCCGGAACATCACTCCAGATGAGGGCTTCTTG	360	
Qy	361	AGTGTGGGTCAAAATAGTTTTCATGTTCCCAAAATGGCCAAAACCTGACAGTTTAAACGCT	420	
Db	361	AGTGTGGGTCAAAATAGTTTTCATGTTCCCAAAATGGCCAAAACCTGACAGTTTAAACGCT	420	
Qy	421	GTCTTGGAACTTAATATGACAAAAGCGTGATCTCATCCAAAGATGAATCAAGTTTGTTTCG	480	
Db	421	GTCTTGGAACTTAATATGACAAAAGCGTGATCTCATCCAAAGATGAATCAAGTTTGTTTCG	480	
Qy	481	TTGAAATGCTAACCGCCAGTTGGTCAAAAGAAACTTTCAAAAAGTCGCCATACCGTTTGT	540	
Db	481	TTGAAATGCTAACCGCCAGTTGGTCAAAAGAAACTTTCAAAAAGTCGCCATACCGTTTGT	540	

Qy	541	CTTGTGCTGATTGATTGATGACGAATGCTCAAAATAATCTCATTATGCTTTAGCGAGTCT	600
Db	541	CTTGTGCTGATTGATTGATGACGAATGCTCAAAATAATCTCATTATGCTTTAGCGAGTCT	500
Qy	601	CTCTATCGCTTCTGAAACCCCGGTGCACCTGTGCCGAAACGCAAAATGGGAAACACCCGT	660
Db	601	CTCTATCGCTTCTGAAACCCCGGTGCACCTGTGCCGAAACGCAAAATGGGAAACACCCGT	560
Qy	661	TTTTTGGATGATATGCAATGTCTCCACATGTATGCTTCCAGATCTCGTGGGATATCT	720
Db	661	TTTTTGGATGATATGCAATGTCTCCACATGTATGCTTCCAGATCTCGTGGGATATCT	720
Qy	721	GCTGATAGCTTAACGTTTCATGATCAAAAATTAACTCTTCTAAACCCCTACTTGCACGAAT	780
Db	721	GCTGATAGCTTAACGTTTCATGATCAAAAATTAACTCTTCTAAACCCCTACTTGCACGAAT	780
Qy	781	ATATAACAGAGGAAGCTGCCCTGTCTTAAACCTTTTTTTTATCATCTATTATAGCTT	840
Db	781	ATATAACAGAGGAAGCTGCCCTGTCTTAAACCTTTTTTTTATCATCTATTATAGCTT	840
Qy	841	ACTTTCATTAATTCGCACTGCTTCCAAATTCGACAGCTTTTGAATTTAAACGACTTTTAAACGA	900
Db	841	ACTTTCATTAATTCGCACTGCTTCCAAATTCGACAGCTTTTGAATTTAAACGACTTTTAAACGA	900
Qy	901	CAACTTGGAGAAGATCAAAAAACAACCTAAATTTATTCGAAGGATCCAAACGATGAGATTTCCT	960
Db	901	CAACTTGGAGAAGATCAAAAAACAACCTAAATTTATTCGAAGGATCCAAACGATGAGATTTCCT	960
Qy	961	TCAATTTTACTGCAGTTTTATTCGAGCATCTCCGCACTAGCTGCTCCAGTCAACACT	1020
Db	961	TCAATTTTACTGCAGTTTTATTCGAGCATCTCCGCACTAGCTGCTCCAGTCAACACT	1020
Qy	1021	ACAACAGAAGATGAACCGGCACAAATTCGGCTGAAGCTGTCTATCGTTACTCAGATTTA	1080
Db	1021	ACAACAGAAGATGAACCGGCACAAATTCGGCTGAAGCTGTCTATCGTTACTCAGATTTA	1080
Qy	1081	GAAGGGATTTCGATGTTGCTGTTTTGCGCAATTTCCACAGCACAATAACGGGTTATTG	1140
Db	1081	GAAGGGATTTCGATGTTGCTGTTTTGCGCAATTTCCACAGCACAATAACGGGTTATTG	1140
Qy	1141	TTTTATAATCTACTATTGTCAGCATTTGCTTAAAGAGAAAGGGTATCTCTCGAGAAA	1200
Db	1141	TTTTATAATCTACTATTGTCAGCATTTGCTTAAAGAGAAAGGGTATCTCTCGAGAAA	1200
Qy	1201	AGAGAGCTGAAGCCAGGCCAGGCCAGGCCAGGCCAGGAAACAATGTATCATG	1260
Db	1201	AGAGAGCTGAAGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGAAACAATGTATCATG	1260
Qy	1261	GCCCCCGGAGAGGATTAATTTGGCTTCCCGGTTCTACCGCCAGCAGTGCACGGAG	1320
Db	1261	GCCCCCGGAGAGGATTAATTTGGCTTCCCGGTTCTACCGCCAGCAGTGCACGGAG	1320
Qy	1321	AGAGTGTCTGTTTGTATGACAGTGTCCGGGATTCOCGTGTTCCACCCCAAGGCC	1380
Db	1321	AGAGTGTCTGTTTGTATGACAGTGTCCGGGATTCOCGTGTTCCACCCCAAGGCC	1380
Qy	1381	ATCGAGAACACTCAAGAGAGAAATGTCCCTTCTAAGTGTGGCGTAGAATTCCTTAGG	1440
Db	1381	ATCGAGAACACTCAAGAGAGAAATGTCCCTTCTAAGTGTGGCGTAGAATTCCTTAGG	1440
Qy	1441	CGGCCGGAATTAATTCGCTTACAGATGACTGTTCTCAGTTCAAGTGGGCACATTACG	1500
Db	1441	CGGCCGGAATTAATTCGCTTACAGATGACTGTTCTCAGTTCAAGTGGGCACATTACG	1500
Qy	1501	AGAAGACCGGTCTTGCTAGATTCTTAATCAAGAGATGTCAGATGCCATTGTCCTGAGAG	1560
Db	1501	AGAAGACCGGTCTTGCTAGATTCTTAATCAAGAGATGTCAGATGCCATTGTCCTGAGAG	1560
Qy	1561	ATGAGGCTTCAATTTTGTACTTTTTTTTATTTGTAACCTATATAGTATAGATTTTTTTT	1620
Db	1561	ATGAGGCTTCAATTTTGTACTTTTTTTTATTTGTAACCTATATAGTATAGATTTTTTTT	1620
Qy	1621	GTCATTTTGTGTTCTTCTCGTACGAGCTTGCTCTCTGATCAGCCTATCTCGCAGCTGATCAA	1680
Db	1621	GTCATTTTGTGTTCTTCTCGTACGAGCTTGCTCTCTGATCAGCCTATCTCGCAGCTGATCAA	1680

1415 GTCAATTTGTTCTTCTCGTACGAGCTGCTCTGATCAGCCTATCTCGAGCTGATGAA 1474
1681 TATCTTGTGTAGGGTTTGGGAAATCATTCGAGTTTGATGTTTCTTGGTATTCC 1740
1475 TATCTTGTGTAGGGTTTGGGAAATCATTCGAGTTTGATGTTTCTTGGTATTCC 1534
1741 ACTCTCTTTCAGAGTACAGAGATTAAGTGAAGAAGTTTCGTTTGTGCAAGCTTATCGATAA 1800
1535 ACTCTCTTTCAGAGTACAGAGATTAAGTGAAGAAGTTTCGTTTGTGCAAGCTTATCGATAA 1594
1801 GCTTTAATGCGGTAGTTTATCACAGTTAAATGTCTAAGCGAGTCAGGACCGGTGATGAA 1860
1595 GCTTTAATGCGGTAGTTTATCACAGTTAAATGTCTAAGCGAGTCAGGACCGGTGATGAA 1654
1861 ATCTAAACAATGCGTCAATCGTCATCTCGGACCGCTCACCTGGATGCTGTAGGCATAGG 1920
1655 ATCTAAACAATGCGTCAATCGTCATCTCGGACCGCTCACCTGGATGCTGTAGGCATAGG 1714
1921 CTGTGTTATGCGGTATCTGCGGGGCTCTTTCGCGGATATGCTCAATTCGACAGCATCGC 1980
1715 CTGTGTTATGCGGTATCTGCGGGGCTCTTTCGCGGATATGCTCAATTCGACAGCATCGC 1774
1981 CAGTCACATATGCGGTATCTGCGGGGCTCTTTCGCGGATATGCTCAATTCGACAGCATCGC 2040
1775 CAGTCACATATGCGGTATCTGCGGGGCTCTTTCGCGGATATGCTCAATTCGACAGCATCGC 1834
2041 TCTCGGACATGTCGACCGCTTTCGCGGCGCCGAGTCTGCTCGCTTTCGCTTTCGCTTTCG 2100
1835 TCTCGGACATGTCGACCGCTTTCGCGGCGCCGAGTCTGCTCGCTTTCGCTTTCGCTTTCG 1894
2101 AGCCACTATGACATACGCGATCTAGCGGACGACACCGCTCTGCGATCTATCGATCTA 2160
1895 AGCCACTATGACATACGCGATCTAGCGGACGACACCGCTCTGCGATCTATCGATCTA 1954
2161 AATGTAAGTTAAATCTCTAAATAATTAATTAAGTCCAGTTTCTCCATACGAACCTTAA 2220
1955 AATGTAAGTTAAATCTCTAAATAATTAATTAAGTCCAGTTTCTCCATACGAACCTTAA 2014
2221 CAGCATTCGCGTACATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAA 2280
2015 CAGCATTCGCGTACATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAA 2074
2281 TATGTTTCAGTCCCTCAGGAGTACGCTTGTGAAGTATGATGAAGTCTGGAAGTTCGAG 2340
2075 TATGTTTCAGTCCCTCAGGAGTACGCTTGTGAAGTATGATGAAGTCTGGAAGTTCGAG 2134
2341 TGTAACTCCGCTGATTTGACGGGCATATCGGTACGTTGGCAAGTGTGTTGTTACCGG 2400
2135 TGTAACTCCGCTGATTTGACGGGCATATCGGTACGTTGGCAAGTGTGTTGTTACCGG 2194
2401 AGGAGTATCTCCAACTCTCTGAGAGTAGGCAACCAACACAGATCCAGCGTGT 2460
2195 AGGAGTATCTCCAACTCTCTGAGAGTAGGCAACCAACACAGATCCAGCGTGT 2254
2461 GTACTTGATCAACATAAGAGAAGCATCTCGATTTGCGAGATCAAGTGTTCAGGAGCGT 2520
2255 GTACTTGATCAACATAAGAGAAGCATCTCGATTTGCGAGATCAAGTGTTCAGGAGCGT 2314
2521 ACTGATTCGATTCGAAAGCTCTGCTAGGTTGCAACCGTAGGTTGTAGAGTGTG 2580
2315 ACTGATTCGATTCGAAAGCTCTGCTAGGTTGCAACCGTAGGTTGTAGAGTGTG 2374
2581 CAATACACTTTCGATCAATTTCAACCCCTTGGCAACTGCAAGCTTGGTGTGGAACAGCAT 2640
2375 CAATACACTTTCGATCAATTTCAACCCCTTGGCAACTGCAAGCTTGGTGTGGAACAGCAT 2434
2641 CTTCAATTCGGAAGCTCTTGTCTGTCATATCGACGCGCAAGATCACCTGGGAAT 2700
2435 CTTCAATTCGGAAGCTCTTGTCTGTCATATCGACGCGCAAGATCACCTGGGAAT 2494
2701 CAATACCAATGTTGAGCTTGAGACAGAGGCTCTGAGGCAAGAAATCTGATCAGCGTATT 2760

2495 CAATACCAATGTTACGCTTGAGACAGAGGCTGTAGGCAACGAAATCTGGATCAGCGTATT 2554
2761 TATCAGCAATAACTAGAACTTTCAGAAAGCCCGAGGAGCATGTCAATACTACACAGGCGTG 2820
2555 TATCAGCAATAACTAGAACTTTCAGAAAGCCCGAGGAGCATGTCAATACTACACAGGCGTG 2614
2821 ATGTGTCAATTTTGAACCAATCATCTTTGGCAGCAGTAAGCAACTGTTTCTCGGACCAATA 2880
2615 ATGTGTCAATTTTGAACCAATCATCTTTGGCAGCAGTAAGCAACTGTTTCTCGGACCAATA 2674
2881 TTTTGTCACTATAGAAACAGTTTCTGTTCCGTAGCCATAGCAGCTACTGCTGGGCGC 2940
2675 TTTTGTCACTATAGAAACAGTTTCTGTTCCGTAGCCATAGCAGCTACTGCTGGGCGC 2734
2941 CTCTCTCTAGCAGATACACTTTAGCACCAACCTTTGTGGGCAACGTFAGATGACTTCTGGGG 3000
2735 CTCTCTCTAGCAGATACACTTTAGCACCAACCTTTGTGGGCAACGTFAGATGACTTCTGGGG 2794
3001 TAAGGTTACCATCTCTTCTTAGTGAGATGCAAAACAAATTTCTTTGCAACCGCAACTT 3060
2795 TAAGGTTACCATCTCTTCTTAGTGAGATGCAAAACAAATTTCTTTGCAACCGCAACTT 2854
3061 TGGCAGGAAACACCCAGCATCAGGGAAGTGGAAAGGAGCAATTTGCGGTTCCACCAGGAATAT 3120
2855 TGGCAGGAAACACCCAGCATCAGGGAAGTGGAAAGGAGCAATTTGCGGTTCCACCAGGAATAT 2914
3121 AGAGGCAACTTTCTCAATAGGTTCTTGCAAAACGAGCAGCACTACACAGGCGCAAGTCT 3180
2915 AGAGGCAACTTTCTCAATAGGTTCTTGCAAAACGAGCAGCACTACACAGGCGCAAGTCT 2974
3181 CAACCTTGCAACGCTCTCCGTTAGTTAGCTTCAATGGAATTTCTGCAACGTTATCTATAGAGA 3240
2975 CAACCTTGCAACGCTCTCCGTTAGTTAGCTTCAATGGAATTTCTGCAACGTTATCTATAGAGA 3034
3241 GATCAATGGCTCTTAAACGTTATCTGCAATTTGCAATTAAGTTCTCTGGGAAAGAGGCTT 3300
3035 GATCAATGGCTCTTAAACGTTATCTGCAATTTGCAATTAAGTTCTCTGGGAAAGAGGCTT 3094
3301 CTAACACAGGTTCTTCAAAAGCAGCTCCATCAAACTTCGCGAGTTAGTTCTTAAAGGCGTT 3360
3095 CTAACACAGGTTCTTCAAAAGCAGCTCCATCAAACTTCGCGAGTTAGTTCTTAAAGGCGTT 3154
3361 TGTCAACATTTTGACGAAATTTGTCGAAATTTGTTTGAATTTCCATAATCTGTTCCG 3420
3155 TGTCAACATTTTGACGAAATTTGTCGAAATTTGTTTGAATTTCCATAATCTGTTCCG 3214
3421 TTTTCTGATAGGACGAGAGGCGATCTTCAATTTCTTGTGAGGAGCGCTTAGAAGCT 3480
3215 TTTTCTGATAGGACGAGAGGCGATCTTCAATTTCTTGTGAGGAGCGCTTAGAAGCT 3274
3481 CAATTTTTCGCAAAATTTCAATAGCCTTTCAGAAAGGAGCTTCTTTAGGTTTGAATTTCTTCTT 3540
3275 CAATTTTTCGCAAAATTTCAATAGCCTTTCAGAAAGGAGCTTCTTTAGGTTTGAATTTCTTCTT 3334
3541 TAGGTTGTTCTTGTGTTATCTGCTGGCTGGATCTTCTTCTTCTTCTAGTACCTTAGGG 3600
3335 TAGGTTGTTCTTGTGTTATCTGCTGGCTGGATCTTCTTCTTCTTCTAGTACCTTAGGG 3394
3601 ACTTCATATCCAGGTTTCTCTCCACCTCGTCCACGTCACACCGTACTTGGCAGCATCTAA 3660
3395 ACTTCATATCCAGGTTTCTCTCCACCTCGTCCACGTCACACCGTACTTGGCAGCATCTAA 3454
3661 CTAATGCAAAATAAATAAAGTCAGCAATTCGCGAGGCTATATCTTCTTGAATTTAGCTT 3720
3455 CTAATGCAAAATAAATAAAGTCAGCAATTCGCGAGGCTATATCTTCTTGAATTTAGCTT 3514
3721 CTGCAAGTTCTACGCTTCTCTCCCTTAAATTTTAGGCTTCAACAAAATTTCTGCTGCAATA 3780
3515 CTGCAAGTTCTACGCTTCTCTCCCTTAAATTTTAGGCTTCAACAAAATTTCTGCTGCAATA 3574
3781 ACCGTTTCTGATTAAGAACCTTCTGAGCAATTTGCTTCTACGATCCCAAGGTCGCTTCCA 3840
3575 ACCGTTTCTGATTAAGAACCTTCTGAGCAATTTGCTTCTACGATCCCAAGGTCGCTTCCA 3634

Qy	3841	TGGCTTAAACACCTTTTGATTGSCCAAAACAGAAAGTGGCTTCCAAGTGACAGAAACCAA	3900
Db	3635	TGGCTTAAACACCTTTTGATTGSCCAAAACAGAAAGTGGCTTCCAAGTGACAGAAACCAA	3694
Qy	3901	CACCTGTTTCTCAACACAAAATTTCAAGCAGTCTCATCACAAATCCAAATTCGATACCCA	3960
Db	3695	CACCTGTTTCTCAACACAAAATTTCAAGCAGTCTCATCACAAATCCAAATTCGATACCCA	3754
Qy	3961	GCAACTTTTTCAGTGTGCTCCAGATGTAGCACCTTTTATACCAAAACCGTGACGACGATTT	4020
Db	3755	GCAACTTTTTCAGTGTGCTCCAGATGTAGCACCTTTTATACCAAAACCGTGACGACGATTT	3814
Qy	4021	GGTAGACTCCAGTTTGTGTCCCTTATAGCCTCCGGAATAGACTTTTGGACGAGTACACCA	4080
Db	3815	GGTAGACTCCAGTTTGTGTCCCTTATAGCCTCCGGAATAGACTTTTGGACGAGTACACCA	3874
Qy	4081	GGCCCAACGAGTAATTAGAAAGATCGACCACCAAAGTAGTGAATAGACCATCGGCGCGGT	4140
Db	3875	GGCCCAACGAGTAATTAGAAAGATCGACCACCAAAGTAGTGAATAGACCATCGGCGCGGT	3934
Qy	4141	CAGTAGTCAAAAGAGCCCAAAATTTCACTGACGAGGAACCTTTTGGACATCTTTCAGAAA	4200
Db	3935	CAGTAGTCAAAAGAGCCCAAAATTTCACTGACGAGGAACCTTTTGGACATCTTTCAGAAA	3994
Qy	4201	GTTCGTATTTCAGTAGTCAATTTGCCGAGCATCAATATATGGGGATTATACCAAGAAGCAACAG	4260
Db	3995	GTTCGTATTTCAGTAGTCAATTTGCCGAGCATCAATATATGGGGATTATACCAAGAAGCAACAG	4054
Qy	4261	TGGAAGTCAACATCTTACCAACTTTTCGGCTCTCAGAAAAAGCATAAACGTTTCTACTACCGC	4320
Db	4055	TGGAAGTCAACATCTTACCAACTTTTCGGCTCTCAGAAAAAGCATAAACGTTTCTACTACCGC	4114
Qy	4321	CATTAGTGAACCTTTTCAAAATCGCCAGTCGGAGAGAAAAGCAGCAGCGATACTAGCAT	4380
Db	4115	CATTAGTGAACCTTTTCAAAATCGCCAGTCGGAGAGAAAAGCAGCGATACTAGCAT	4174
Qy	4381	TAGCGGCAAGGATGCAACTTTTATCAACACGAGGTCTCTATAGATAAACCTAGCGCCTGGGA	4440
Db	4175	TAGCGGCAAGGATGCAACTTTTATCAACACGAGGTCTCTATAGATAAACCTAGCGCCTGGGA	4234
Qy	4441	TCATCTCTTGGACAACTCTTTCTGCCAAATCTTAGTTCCTCAAAATCACTTCATTGATACCAT	4500
Db	4235	TCATCTCTTGGACAACTCTTTCTGCCAAATCTTAGTTCCTCAAAATCACTTCATTGATACCAT	4294
Qy	4501	TATTGTACAACTTGAGCAAGTTGTCGATCAGTCTCTCAAAATGGTCTCTGTAACCGATG	4560
Db	4295	TATTGTACAACTTGAGCAAGTTGTCGATCAGTCTCTCAAAATGGTCTCTGTAACCGATG	4354
Qy	4561	ACTCAACTTGCACATTAACCTTGAAGTCACTCGATTGAGTGAACCTTGATCAGTGTGCA	4620
Db	4355	ACTCAACTTGCACATTAACCTTGAAGTCACTCGATTGAGTGAACCTTGATCAGTGTGCA	4414
Qy	4621	GCTGGTCAGCAGCATAGGGAACACGGCTTTTCTTACCAACTTCAGGANTTATCAAACT	4680
Db	4415	GCTGGTCAGCAGCATAGGGAACACGGCTTTTCTTACCAACTTCAGGANTTATCAAACT	4474
Qy	4681	CTGCACAACTTGCCTATGTCAGGTAGCAAGGGAATGTCTACTTTGAAGTCGGACAGTGAG	4740
Db	4475	CTGCACAACTTGCCTATGTCAGGTAGCAAGGGAATGTCTACTTTGAAGTCGGACAGTGAG	4534
Qy	4741	TGTAGTCTTGAAGAAATTTCTGAAGCCGTAATTTTATTATCATGTAGTCACTCAGGAGA	4800
Db	4535	TGTAGTCTTGAAGAAATTTCTGAAGCCGTAATTTTATTATCATGTAGTCACTCAGGAGA	4594
Qy	4801	TCCTCTACCGCGGACGATCGTGSCCGA-----	4828
Db	4595	TCCTCTACCGCGGACGATCGTGSCCGA-----	4654
Qy	4829	-----	4828
Db	4655	CTCGTGAAGAAGGTGTGCTGACTCATACAGGCGCTGAATCGGCCCATCATCCAGCCAGA	4714

[illegible]

Db	5795	TGATATATTTTTTATCTTTGTGCAATGTAAACATCATCAGAGATTTTGTGACACAAACGTGGCTTTC	5854
Qy	4829	-----CCTGCAGGTTCGGCATCACCGCGGCCACAGGTGCGGTTCCTCGCGCCCATAT	4879
Db	5855	CCCCCCCCCTCGAGGTGGCATCACCGCGGCCACAGGTGCGGTTCCTGGCGCCTATAT	5914
Qy	4880	CGCCGACATCACCGATCGGGGAAGATCGGGTTCGCCACTTCGGGCTCATGAGCGCTGT	4939
Db	5915	CGCCGACATCACCGATCGGGGAAGATCGGGTTCGCCACTTCGGGCTCATGAGCGCTGT	5974
Qy	4940	CGGCGTGGGTATGTGTGCAGCCCGCTGGCGCGGGGACTGTTGGCGGCCATCTCTCTGCA	4999
Db	5975	CGGCGTGGGTATGTGTGCAGCCCGCTGGCGCGGGGACTGTTGGCGGCCATCTCTCTGCA	6034
Qy	5000	TGCACCATCTCTTCGCGCGCGGTGCTCAACGGGCTCAACCTACTACTCTGGGCTGCTTCCT	5059
Db	6035	TGCACCATCTCTTCGCGCGCGGTGCTCAACGGGCTCAACCTACTACTCTGGGCTGCTTCCT	6094
Qy	5060	AATGACGAGTCCGATAAAGGAGAGCGTTCGAGTATCTATGATTCGGAAGTATGGGAATGGT	5119
Db	6095	AATGACGAGTCCGATAAAGGAGAGCGTTCGAGTATCTATGATTCGGAAGTATGGGAATGGT	6154
Qy	5120	GATACCCGATCTTTCAGTGTCTTGAGTCTTCCTATCAGATTATGCCCACTAAAGCAAC	5179
Db	6155	GATACCCGATCTTTCAGTGTCTTGAGTCTTCCTATCAGATTATGCCCACTAAAGCAAC	6214
Qy	5180	CGGAGGAGGATTTTCATGTGTAATTTCTTGACTTTTGGTTCATGATGCACTCGAACTG	5239
Db	6215	CGGAGGAGGATTTTCATGTGTAATTTCTTGACTTTTGGTTCATGATGCACTCGAACTG	6274
Qy	5240	TGAGCATCTCTCGGTATGACAGCAGAAATCTCTCTTCGAGACAGTAATGAAGTCCC	5299
Db	6275	TGAGCATCTCTCGGTATGACAGCAGAAATCTCTCTTCGAGACAGTAATGAAGTCCC	6334
Qy	5300	ACCAATAAAGAAATCTCTTTTATCAGGAACAAATCTCTGTTTCGAACTTTTCGGTGCC	5359
Db	6335	ACCAATAAAGAAATCTCTTTTATCAGGAACAAATCTCTGTTTCGAACTTTTCGGTGCC	6394
Qy	5360	TTGAACHTAAATGTAGAGTGGATATGTCGGGTAGGAATGGAGCGGCAATGCTTACC	5419
Db	6395	TTGAACHTAAATGTAGAGTGGATATGTCGGGTAGGAATGGAGCGGCAATGCTTACC	6454
Qy	5420	TTCTGGACCTTCAAGAGGTATGTAGGGTTTGTAGATACTGATGCCAACTTCAGTGACAAC	5479
Db	6455	TTCTGGACCTTCAAGAGGTATGTAGGGTTTGTAGATACTGATGCCAACTTCAGTGACAAC	6514
Qy	5480	GTGCTATTTCGTTCAAAACCATTCGGAATCCAGAGAAATCAAAGTTGTTGTCTACTATT	5539
Db	6515	GTGCTATTTCGTTCAAAACCATTCGGAATCCAGAGAAATCAAAGTTGTTGTCTACTATT	6574
Qy	5540	GATCCAGCCAGTGGGTCCTTGAACCTCACAATAGTGTGCTCGTGTCTTCAGGTCATCTT	5599
Db	6575	GATCCAGCCAGTGGGTCCTTGAACCTCACAATAGTGTGCTCGTGTCTTCAGGTCATCTT	6634
Qy	5600	TGTATGAATAAATCTAGTCTTTGATCTAAATAATCTTTGACGAGCCAGGCGATAAATACC	5659
Db	6635	TGTATGAATAAATCTAGTCTTTGATCTAAATAATCTTTGACGAGCCAGGCGATAAATACC	6694
Qy	5660	CAATCTAAATCTTTTAAACGTTAAAGGACAAGTATCTCTGCTGTATTAAACCCC	5719
Db	6695	CAATCTAAATCTTTTAAACGTTAAAGGACAAGTATCTCTGCTGTATTAAACCCC	6754
Qy	5720	AAATCAGTCTGATCTCATCTCAATCTTGAGGGGCACTATCTCTGTTTATGAGAAAT	5779
Db	6755	AAATCAGTCTGATCTCATCTCAATCTTGAGGGGCACTATCTCTGTTTATGAGAAAT	6814
Qy	5780	TTGCGAGATCGGATATCGAGAAAGGTACGCTGATTTTAAACGTGAAATTTATCTCAA	5839
Db	6815	TTGCGAGATCGGATATCGAGAAAGGTACGCTGATTTTAAACGTGAAATTTATCTCAA	6874
Qy	5840	GATCTCTCCCTCGCGGCTTTCGGTGATACGGTGAAGAAACCTCTGACATCAGCTCCCG	5899

6875	Db	GATCTCTGCGCTTTCGGGTATGACGGTGAAAACTCTTGACACATGACAGCTCCCG	6933
5900	Qy	GAGACGGTCAAGCTTGTCTGTAAGCGGATGCCGGAGCAGACAAGCCCGTCAAGGCGCG	5959
6935	Db	GAGACGGTCAAGCTTGTCTGTAAGCGGATGCCGGAGCAGACAAGCCCGTCAAGGCGCG	6994
5960	Qy	TCAGCGGTTGTGGCGGGTGTGGGGCGACCATGACCCAGTCACTGATCGATAGCGGA	6019
6995	Db	TCAGCGGTTGTGGCGGGTGTGGGGCGACCATGACCCAGTCACTGATCGATAGCGGA	7054
6020	Qy	GTGTATACTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACATATGC	6079
7055	Db	GTGTATACTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACATATGC	7114
6080	Qy	GGTGTGAAATACCGCACAGATGCGTAAAGAGAAAAATACCGCATCAGGCGCTCTCCGCTT	6139
7115	Db	GGTGTGAAATACCGCACAGATGCGTAAAGAGAAAAATACCGCATCAGGCGCTCTCCGCTT	7174
6140	Qy	CCTCGCTCACTGACTCGCTCGCTCGGTGCTTGGCTGCGCGAGCGGTATCAGCTCACT	6199
7175	Db	CCTCGCTCACTGACTCGCTCGCTCGGTGCTTGGCTGCGCGAGCGGTATCAGCTCACT	7234
6200	Qy	CAAAAGCGGTAAATACGGTTATCCACAGAATCAGGGGATAACGACGAGAAACATGTGAG	6259
7235	Db	CAAAAGCGGTAAATACGGTTATCCACAGAATCAGGGGATAACGACGAGAAACATGTGAG	7294
6260	Qy	CAAAAGCCGCAAAAGCCAGAGAACGTTAAAGGCGCGGTTGCTGGCGTTTTCATATA	6319
7295	Db	CAAAAGCCGCAAAAGCCAGAGAACGTTAAAGGCGCGGTTGCTGGCGTTTTCATATA	7354
6320	Qy	GGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAAGGTGCGGAAACC	6379
7355	Db	GGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAAGGTGCGGAAACC	7414
6380	Qy	CGACAGACTATAAGATACAGCGGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCCCTG	6439
7415	Db	CGACAGACTATAAGATACAGCGGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCCCTG	7474
6440	Qy	TTCCGACCCCTGCGCCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGCGC	6499
7475	Db	TTCCGACCCCTGCGCCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGCGC	7534
6500	Qy	TTTCTCATAGCTACGCTGTAGTATCTCATGTTCCGTTAGGTCGTCGCTCCAAGCTGG	6559
7535	Db	TTTCTCATAGCTACGCTGTAGTATCTCATGTTCCGTTAGGTCGTCGCTCCAAGCTGG	7594
6560	Qy	GCTGTGTGCAGAACCCCGCTTCAGCGCCACCGCTCGCCTTATCCGGTAACTATCGTC	6619
7595	Db	GCTGTGTGCAGAACCCCGCTTCAGCGCCACCGCTCGCCTTATCCGGTAACTATCGTC	7654
6620	Qy	TTGAGTCCAAACCCGGTAAGACAAGCTTATCGCCACTGGCAGACGCCACTGGTAAACGGA	6679
7655	Db	TTGAGTCCAAACCCGGTAAGACAAGCTTATCGCCACTGGCAGACGCCACTGGTAAACGGA	7714
6680	Qy	TTACGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACTACG	6739
7715	Db	TTACGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACTACG	7774
6740	Qy	GCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA	6799
7775	Db	GCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA	7834
6800	Qy	AAAAGTTGTTAGCTCTGTATCCCGCAAAACAAACCCCGCTGGTAGCGGTGGTTTTTTG	6859
7835	Db	AAAAGTTGTTAGCTCTGTATCCCGCAAAACAAACCCCGCTGGTAGCGGTGGTTTTTTG	7894
6860	Qy	TTTGAAGCAGCAGATTACCGCGCAGAAAAAAGGATCTCAAGAGATCTCTTTGATCTTTT	6919
7895	Db	TTTGAAGCAGCAGATTACCGCGCAGAAAAAAGGATCTCAAGAGATCTCTTTGATCTTTT	7954
6920	Qy	CTACGGGCTCTGACGCTCAGTGGAAACGAAACTCACTTAAGGGATTTTGGTCAATGAGAT	6979
7955	Db	CTACGGGCTCTGACGCTCAGTGGAAACGAAACTCACTTAAGGGATTTTGGTCAATGAGAT	8014

482 QY TGAATGCTAAACGGCCAGTTGGTCAAAAGAACTTCCAAAGATCGCCATACCGTTTGTG 541
 481 Db TGAATGCTAAACGGCCAGTTGGTCAAAAGAACTTCCAAAGATCGCCATACCGTTTGTG 540
 542 QY TTGTTGGTATTGATGACCAATGCTCAAAATAATCTCAATTAATGCTTGGGCGAGTCTC 601
 541 Db TTGTTGGTATTGATGACCAATGCTCAAAATAATCTCAATTAATGCTTGGGCGAGTCTC 600
 602 QY TCTATCGCTTCTGAACCCCGGTGACCTGTCGGAACGCAATGGGGAACACCCGCTT 661
 601 Db TCTATCGCTTCTGAACCCCGGTGACCTGTCGGAACGCAATGGGGAACACCCGCTT 660
 662 QY TTTGGATGATTATGATGCTCCACATTTGTAATGCTTCCAGATTCGGTGGGAATACCTG 721
 661 Db TTTGGATGATTATGATGCTCCACATTTGTAATGCTTCCAGATTCGGTGGGAATACCTG 720
 722 QY CTGATAGCTTAACGTTCAATGATCAAAATTTAACTGTTCTAACCCCTACTTTGACAGCAATA 781
 721 Db CTGATAGCTTAACGTTCAATGATCAAAATTTAACTGTTCTAACCCCTACTTTGACAGCAATA 780
 782 QY TATAAACAAGAGGAGTGCCTGCTTAAACCTTTTATTAATCATCATATATAGCTTA 841
 781 Db TATAAACAAGAGGAGTGCCTGCTTAAACCTTTTATTAATCATCATATATAGCTTA 840
 842 QY CTTTTCATTAATGCGACTGGTTCCTAATGACAGCTTTTGTGATTTTAAACGACTTTTAAACGAC 901
 841 Db CTTTTCATTAATGCGACTGGTTCCTAATGACAGCTTTTGTGATTTTAAACGACTTTTAAACGAC 900
 902 QY AACTTGAGAAGATCAAAACCAACTTAATTTCAAGAGATCCAAACGATGAGATTTTCCTT 961
 901 Db AACTTGAGAAGATCAAAACCAACTTAATTTCAAGAGATCCAAACGATGAGATTTTCCTT 960
 962 QY CAATTTTACTGCGAGTTTATTCGAGCATCTCCGATGATGCTCCAGTCAACACTA 1021
 961 Db CAATTTTACTGCGAGTTTATTCGAGCATCTCCGATGATGCTCCAGTCAACACTA 1020
 1022 QY CAACAGAAGATGAACGGCAAAATTCGGCTGAAGCTGTGATCGGTTACTCGAGATTAG 1081
 1021 Db CAACAGAAGATGAACGGCAAAATTCGGCTGAAGCTGTGATCGGTTACTCGAGATTAG 1080
 1082 QY AAGGGGATTCGATGCTGCTTTGCGCATTTTCAACAGACCAAAATAACGGTTATTGT 1141
 1081 Db AAGGGGATTCGATGCTGCTTTGCGCATTTTCAACAGACCAAAATAACGGTTATTGT 1140
 1142 QY TTATAAATACTACTATTCGAGCATTTGCTGCTAAAGAGAGGGGTATCTCTCGAGAAAA 1201
 1141 Db TTATAAATACTACTATTCGAGCATTTGCTGCTAAAGAGAGGGGTATCTCTCGAGAAAA 1200
 1202 QY GAGAGGCTGAAGCCGAGG----- 1219
 1201 Db GATCGGACCGGGCAGGGGTTGGGGAAGAGAGGACCCCAAAAGCTGACCCCTTTAG 1260
 1220 QY -----CCAGGCCAGGCCAGGCCAG 1242
 1261 Db CCTACAAGAGTTTATCCCAATGTTGGCCGAGAGACCTTAGGCCGAGGAGGTATG 1320
 1243 QY GAAGAAACAT----- 1252
 1321 Db AAGGAGATCTCCAGAAATCCGAGCATTTAAGAACTCACCCCAATTAACACCCCG 1380
 1253 QY -----GTATCATGCCCGCCGAGAGGATAAATTTGGTTCCTCCCGGTGTC 1299
 1381 Db ACATCATATTTAAGGATGAAGAAAAACCGGAGCCGACAGGCTGATGACTCAGAGGTGTA 1440
 1300 QY AC-----CGCCAGCAGTGCACGGAGAGAGGTTCCTGTTT 1334
 1441 Db AGGACAAGTTGAACGCTTTGGCCATCTCGGTGATGAACAGTGGCCAGGAGTGAACCTGC 1500
 1335 QY TGATGACAGTTCGGGGGATTTCCCGTGG----- 1362
 1501 Db GGGTGACCGAGGGCTGGGACGAAGATGGCCACCACTCAGAGGAGTCTCTGCACACGAGG 1560

1363 QY ----- 1362
 1561 Db GCCGGCAGTGGACATCACACAGTCTGACCGGACCGCAGCAAGTACGGCATGCTGCCCC 1620
 1363 QY -----TGCTTCCACCCCATGGCCATCGAGAACACCTCAAGAGAAGAAT 1405
 1621 Db GCCTGGCGGTGGAGCGCGGCTTCGACTGGGTGTACTAGAGTCCAAAGGCAATATCCACT 1680
 1406 QY GTCCCTTCTAATA-----GTGGGTAGAAATCCCTAGG-----GCGGCCCGGA 1449
 1681 Db GCTCGGTGAAGAGCAGAGAACTCGGTGGGGCCAAATCGGAGGCTGATTCGGGCCCGGA 1740
 1450 QY ATTAATTCGCTTATGACATGACTCTCTCAGTTTCAAGTTGGGCACTTTACGAGAAGACCG 1509
 1741 Db ATTAATTCGCTTATGACATGACTCTCTCAGTTTCAAGTTGGGCACTTTACGAGAAGACCG 1800
 1510 QY GTCTTGTCTAGATTTCTAATCAAGAGATGTCAGATGCCATTTGCTCGAGATGACAGCT 1569
 1801 Db GTCTTGTCTAGATTTCTAATCAAGAGATGTCAGATGCCATTTGCTCGAGATGACAGCT 1860
 1570 QY TCAATTTTGTACTATTTTATTTTAACTATATAGTATAGGATTTTGTCTCATTTG 1629
 1861 Db TCAATTTTGTACTATTTTATTTTAACTATATAGTATAGGATTTTGTCTCATTTG 1920
 1630 QY TTTCTTCTCGTACGAGCTTGTCTCTGATCAGCTTCTCGAGCTGATGATATATCTTGTG 1689
 1921 Db TTTCTTCTCGTACGAGCTTGTCTCTGATCAGCTTCTCGAGCTGATGATATATCTTGTG 1980
 1690 QY GTAGGGTTTGGGAAATCAATTCGAGTTTGTATGTTTTCTTGGTATTTCCCACTCTCTT 1749
 1981 Db GTAGGGTTTGGGAAATCAATTCGAGTTTGTATGTTTTCTTGGTATTTCCCACTCTCTT 2040
 1750 QY CAGAGTACAGAAGATTAAGTGAGAGTTTCGTTTGTGCAAGCTTATCGATAAGCTTTAATG 1809
 2041 Db CAGAGTACAGAAGATTAAGTGAGAGTTTCGTTTGTGCAAGCTTATCGATAAGCTTTAATG 2100
 1810 QY CGGTAGTTTACACAGTTTAAATTTGCTAACGAGTCCAGCACCGCTGATGAATCTAACA 1869
 2101 Db CGGTAGTTTATCACAGTTTAAATTTGCTAACGAGTCCAGCACCGCTGATGAATCTAACA 2160
 1870 QY TCGGCTCATGCTATCTTCGGCACCGTCACTTGTGATCTGTAGCATPAGGCTTGTAT 1929
 2161 Db TCGGCTCATGCTATCTTCGGCACCGTCACTTGTGATCTGTAGCATPAGGCTTGTAT 2220
 1930 QY GCGGTACTGCGCGGCGCTTTCGGGGATATCGTCCATTCGACAGCATCGCCAGTCACTA 1989
 2221 Db GCGGTACTGCGCGGCGCTTTCGGGGATATCGTCCATTCGACAGCATCGCCAGTCACTA 2280
 1990 QY TCGCGTGTCTAGCGCTATATGCTTGTATGCAATTTCTATCGCACCCGTTCTCGGAGC 2049
 2281 Db TCGCGTGTCTAGCGCTATATGCTTGTATGCAATTTCTATCGCACCCGTTCTCGGAGC 2340
 2050 QY ACTGTCCGACCGGTTTGGCGCGCGCTTCTGCTCGCTTCTGCTTGGAGCCACTAT 2109
 2341 Db ACTGTCCGACCGGTTTGGCGCGCGCTTCTGCTCGCTTCTGCTTGGAGCCACTAT 2400
 2110 QY CGACTACGCGATCATGGCGACCAACCCGCTCTGCTGGATCTATCGAATCTAAATGTAAGT 2169
 2401 Db CGACTACGCGATCATGGCGACCAACCCGCTCTGCTGGATCTATCGAATCTAAATGTAAGT 2460
 2170 QY TAAATCTCTAAATAATTAATAAGTCCAGTTTCTCATACGAACTTACAGCAATTCG 2229
 2461 Db TAAATCTCTAAATAATTAATAAGTCCAGTTTCTCATACGAACTTACAGCAATTCG 2520
 2230 QY GGTGACATCTAGACTTCAACAGAGCCAGATCCATCACTCTTGGGCAATATCTTCA 2289
 2521 Db GGTGACATCTAGACTTCAACAGAGCCAGATCCATCACTCTTGGGCAATATCTTCA 2580
 2290 QY GTCCCTCAGGAGTTACGCTTTGTGAAGTGAATGAATCTTGTGAAGGTTGAGTTAATCTC 2349
 2581 Db GTCCCTCAGGAGTTACGCTTTGTGAAGTGAATGAATCTTGTGAAGGTTGAGTTAATCTC 2640
 2350 QY CGCTGTATTGACGGGCATATCGTACGTTGGCAAGTGGTGGTGGTACCGGAGGAGTAAT 2409

Db 2641 CCGCTGATTTAGCGGCAATATCCGTACGTTGGCAAGTGTGGTACCGGAGAGTAAT 2700
Qy 2410 CTCACAACTCTCTGAGAGTAGGCAACCAACAAACACAGATCCAGCGGTGTGTACTTGAAT 2469
Db 2701 CTCACAACTCTCTGAGAGTAGGCAACCAACAAACACAGATCCAGCGGTGTGTACTTGAAT 2760
Qy 2470 CAAATATAGAGAGCAATCTCGATTTTCAGAGATCAAGTGTTCAGGAGCGTACTGATTTGG 2529
Db 2761 CAAATATAGAGAGCAATCTCGATTTTCAGAGATCAAGTGTTCAGGAGCGTACTGATTTGG 2820
Qy 2530 ACAATTTCAAAGCCCTGCTAGTGTTCGAAACCGATAGGTTGTAGAGTGTGCAATACACT 2589
Db 2821 ACAATTTCAAAGCCCTGCTAGTGTTCGAAACCGATAGGTTGTAGAGTGTGCAATACACT 2880
Qy 2590 TCGGTACAAATTTCAACCCCTTGGCAACTGCAAGCTTGTGTGAAACAGCACTCTCAATTC 2649
Db 2881 TCGGTACAAATTTCAACCCCTTGGCAACTGCAAGCTTGTGTGAAACAGCACTCTCAATTC 2940
Qy 2650 TGGCAAGCTCTTGTCTGTCTATATCGACAGCCCAACAGAAATCACTGGGAATCAATACCAT 2709
Db 2941 TGGCAAGCTCTTGTCTGTCTATATCGACAGCCCAACAGAAATCACTGGGAATCAATACCAT 3000
Qy 2710 GTTCAGTTGAGACAGAGGTTGAGGCAACGAAATCTGATTCAGCGTATTTATCAGCAA 2769
Db 3001 GTTCAGTTGAG- CAGAAAGGTTGAGGCAACGAAATCTGATTCAGCGTATTTATCAGCAA 3059
Qy 2770 TAACTAGAACTTCAGAAAGCCAGCAGCATGTCAATCTGATTCAGCGTCTGATGTGCAT 2829
Db 3060 TAACTAGAACTTCAGAAAGCCAGCAGCATGTCAATCTGATTCAGCGTCTGATGTGCAT 3119
Qy 2830 TTTGAACCATCATCTTGGCAGCAGTAACGAACTGGTTTCTGGACCAAAATATTTGTGCAC 2889
Db 3120 TTTGAACCATCATCTTGGCAGCAGTAACGAACTGGTTTCTGGACCAAAATATTTGTGCAC 3179
Qy 2890 ACTTAGGACAGTTTCTGTTCGTAAGCATAGCAGTACTGCTGGGGCGCTCTGCTA 2949
Db 3180 ACTTAGGACAGTTTCTGTTCGTAAGCATAGCAGTACTGCTGGGGCGCTCTGCTA 3239
Qy 2950 GCACATACACTTAGCACCAACTTGTGGGCAACGTAGATCACTTCTGGGTAAGGGTAC 3009
Db 3240 GCACATACACTTAGCACCAACTTGTGGGCAACGTAGATCACTTCTGGGTAAGGGTAC 3299
Qy 3010 CATCTCTTAGTGGAGATGCAAAAACAAATTTCTTTCGCAACCGAACTTTGGCAGGAA 3069
Db 3300 CATCTCTTAGTGGAGATGCAAAAACAAATTTCTTTCGCAACCGAACTTTGGCAGGAA 3359
Qy 3070 CACCCAGCATCAGGGAAGTGGAGGCAAGTTGCGGTTCCACCGAAATATAGAGGCCAA 3129
Db 3360 CACCCAGCATCAGGGAAGTGGAGGCAAGTTGCGGTTCCACCGAAATATAGAGGCCAA 3419
Qy 3130 CTTTCTCAATAGGTTCTGCAAAACGAGAGCAGACTACACCGGGCAAGTCTCAACTTGCAC 3189
Db 3420 CTTTCTCAATAGGTTCTGCAAAACGAGAGCAGACTACACCGGGCAAGTCTCAACTTGCAC 3479
Qy 3190 AGCTCTCGGTAGTTAGCTTCATCGAATTTCTGACGTTTCTATAGAGAGTCAATGG 3249
Db 3480 AGCTCTCGGTAGTTAGCTTCATCGAATTTCTGACGTTTCTATAGAGAGTCAATGG 3539
Qy 3250 CTCTCTTAACTGTTATCTGCAATTCGATAAGTTCTCTGGGAAAGGAGCTTCTAACACAG 3309
Db 3540 CTCTCTTAACTGTTATCTGCAATTCGATAAGTTCTCTGGGAAAGGAGCTTCTAACACAG 3599
Qy 3310 GTGCTCTTAAGCGCATCTCAAACTTGGCGAGTTAGTCTTAAAGGGCTTTGTCAACCAT 3369
Db 3600 GTGCTCTTAAGCGCATCTCAAACTTGGCGAGTTAGTCTTAAAGGGCTTTGTCAACCAT 3659
Qy 3370 TTTGACGAACTTGTGCAAAATTTGTGAACTTGTGCAATTTCTGTTCCGTTTCTGGA 3429
Db 3660 TTTGACGAACTTGTGCAAAATTTGTGAACTTGTGCAATTTCTGTTCCGTTTCTGGA 3719
Qy 3430 TAGGACGAGAGGCACTTCAATTTCTTGTAGGAGCGCTTAGAAACGTCATTTTGC 3489

Db 3720 TAGGACGAGAGGCACTCTTCAATTTCTTGTAGGAGCGCTTAGAAACGTCATTTTGC 3779
Qy 3490 ACAATTCAAATACGACCTTCAGAAAGGACCTTCTTTAGTTTGGATTCTTCTTTAGGTTGTT 3549
Db 3780 ACAATTCAAATACGACCTTCAGAAAGGACCTTCTTTAGTTTGGATTCTTCTTTAGGTTGTT 3839
Qy 3550 CTTTGGGTATCTCTGCTTGGCATCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3609
Db 3840 CTTTGGGTATCTCTGCTTGGCATCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3899
Qy 3610 CCAGGTTTCTCTCCACTCTCTCAACGTCACACCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3669
Db 3900 CCAGGTTTCTCTCCACTCTCTCAACGTCACACCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3959
Qy 3670 AATAAATAGTCAGCAGCAATTCAGGCTATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3729
Db 3960 AATAAATAGTCAGCAGCAATTCAGGCTATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4019
Qy 3730 CATCAGCTTCT 3788
Db 4020 CATCAGCTTCT 4079
Qy 3789 GTATAGAACTTCTGAGCATTTGCTCTTACGATCCCAAGGTCGCTTCCATGGCTCTA 3848
Db 4080 GTATAGAACTTCTGAGCATTTGCTCTTACGATCCCAAGGT- GCTTCCATGGCTCTA 4138
Qy 3849 AGACCCCTTTGATTGGCCAAAACAGGAAGTGGCTTCCAAAGTGACAGAAAACCAACACCTGTT 3908
Db 4139 AGACCCCTTTGATTGGCCAAAACAGGAAGTGGCTTCCAAAGTGACAGAAAACCAACACCTGTT 4198
Qy 3909 TGTTCACCAAAATTTCAAGAGTCTCCATCACAATTCGAATCCAGCAACTTT 3968
Db 4199 TGTTCACCAAAATTTCAAGAGTCTCCATCACAATTCGAATCCAGCAACTTT 4258
Qy 3969 TGAAGTCTCCAGCATGACACCTTTATACCAACCGTGACGAGAGATTTGGTAGACT 4028
Db 4259 TGAAGTCTCCAGCATGACACCTTTATACCAACCGTGACGAGAGATTTGGTAGACT 4318
Qy 4029 CCAGTTTGTCTCTTAGCTCCGGAATAGACTTTTGGACGAGTACACAGGCCCAAC 4088
Db 4319 CCAGTTTGTCTCTTAGCTCCGGAATAGACTTTTGGACGAGTACACAGGCCCAAC 4378
Qy 4089 GAGTAATAGAGAGTACAGCCCAAGTAGTGAATAGACCATCCGGGGGGTTCAGTAGTC 4148
Db 4379 GAGTAATAGAGAGTACAGCCCAAGTAGTGAATAGACCATCCGGGGGGTTCAGTAGTC 4438
Qy 4149 AAAGACGCCAAACAAATTTCACTGACAGGCACTTTTTCACATCTTCAGAAAGTTCGTAT 4208
Db 4439 AAAGACGCCAAACAAATTTCACTGACAGGCACTTTTTCACATCTTCAGAAAGTTCGTAT 4498
Qy 4209 TCAGTAGTCAATTTGCCGAGCATCAATTAATGGGGATTTATACAGAGCAACAGTGGAGTTC 4268
Db 4499 TCAGTAGTCAATTTGCCGAGCATCAATTAATGGGGATTTATACAGAGCAACAGTGGAGTTC 4558
Qy 4269 ACATCTACCAACTTTCGGGTCTCAGAAAAGGATTAACAGTTCTACTTACCGCCATTAGTG 4328
Db 4559 ACATCTACCAACTTTCGGGTCTCAGAAAAGGATTAACAGTTCTACTTACCGCCATTAGTG 4618
Qy 4329 AAACTTTTCAAATTCGCCAGTGGAGAGAAAAGGCAACGCGATACTAGCATTTAGCGGC 4388
Db 4619 AAACTTTTCAAATTCGCCAGTGGAGAGAAAAGGCAACGCGATACTAGCATTTAGCGGC 4678
Qy 4389 AAGGATGCAACTTTATCAACCGGCTCTATAGATAACCTAGCGCTGGATCATCCCT 4448
Db 4679 AAGGATGCAACTTTATCAACCGGCTCTATAGATAACCTAGCGCTGGATCATCCCT 4738
Qy 4449 TGGACAACTCTTCTGCCAAATCTAGGTCAAAATCACTTCAATGATGATCACTTATTGTAC 4508
Db 4739 TGGACAACTCTTCTGCCAAATCTAGGTCAAAATCACTTCAATGATGATCACTTATTGTAC 4798
Qy 4509 AACTGAGCAAGTTGTGATCAGCTCTCTCAAAATTTGTCTCTGTGTAACGATGATCAACT 4568
Db 4799 AACTGAGCAAGTTGTGATCAGCTCTCTCAAAATTTGTCTCTGTGTAACGATGATCAACT 4858

QY	4569	TGCACTTAACCTTGAAGCTCAGTCGATTTGAGTGAACCTTGTATCAGGTTGTCCAGCTGGTCA	4628	QY	4998	-----	4997
Db	4859	TGCACATTAACCTTGAAGCTCAGTCGATTTGAGTGAACCTTGTATCAGGTTGTCCAGCTGGTCA	4918	Db	5939	CTGTTGAACAACTCTGGAAAGAAATGCATAAGCTTTTGGCCATTCTCACCGGATTCACTCG	5998
QY	4629	GCAGATAGGAAACACAGGCTTTTCCCTACCAAACTCAAGAAATATCAAACTCTGCAACA	4688	QY	4998	-----	4997
Db	4919	GCAGATAGGAAACACAGGCTTTTCCCTACCAAACTCAAGAAATATCAAACTCTGCAACA	4978	Db	5999	TCACCTCATGGTGAATTTCTCACTTGATAACCTTATTTTGTGACGAGGGGAAATTAATAGTT	6058
QY	4689	CTTGGCTATGCAAGTACGAGGAAATGTCTATCTTGAAGTCGGAAGTGTAGTCT	4748	QY	4998	-----	4997
Db	4979	CTTGGCTATGCAAGTACGAGGAAATGTCTATCTTGAAGTCGGAAGTGTAGTCT	5038	Db	6059	GTATTTGATGTTGGACGAGTCGGAATCCGAGACCGATACAGAGATCTTGCCATCCTATGGA	6118
QY	4749	TGAGAAATCTGAAGCCGTATTTTATTTATCAGTGAAGTCAATCAGGAGATCCTCTAC	4808	QY	4998	-----	4997
Db	5039	TGAGAAATCTGAAGCCGTATTTTATTTATCAGTGAAGTCAATCAGGAGATCCTCTAC	5098	Db	6119	ACTGCTCGGTGAGTTTCTCCTTCATTAACAGAAACGCGCTTTTCAAAAATATGGTATG	6178
QY	4809	GCCGACGCAATCGTGGCCGACCTGCAAGTCCGCAATCACCGGCGCCACAGGTGCGGTGCT	4868	QY	4998	-----	4997
Db	5099	GCCGACGCAATCGTGGCCGACCTGCAAGTCCGCAATCACCGGCGCCACAGGTGCGGTGCT	5158	Db	6179	ATAATCCTGATATGAATAAATGCAAGTTTCATTTGATGCTCGATGAGTTTCTTAATCAG	6238
QY	4869	GCGGCTATATCGCGCATCACCGATCGGGAAGATCGGCTCGCCTCTCGGCTCATG	4928	QY	4998	-----	4997
Db	5159	GCGGCTATATCGCGCATCACCGATCGGGAAGATCGGCTCGCCTCTCGGCTCATG	5218	Db	6239	AATCGTTAATTGGTTGTAACACTGGCAGAGCAITACGCTGACTTGACGGGACGCGGCT	6298
QY	4929	AGCGCTTCTTGGGCTGATGTTGGCAGGCCCCGCTGGCCGGGGGACTGTTGGGCGCC	4988	QY	4998	-----	4997
Db	5219	AGCGCTTCTTGGGCTGATGTTGGCAGGCCCCGCTGGCCGGGGGACTGTTGGGCGCC	5278	Db	6299	TTGTTGAATAAATCGAACTTTTGTGTGAGTTGAAGATCAGATCAGGATCTTCCGACAA	6358
QY	4989	ATCTCCTTG-----	4997	QY	4998	-----	4997
Db	5279	ATCTCCTTGACCTGACGGGGGGGGGGGAAACCAACGTTGTCTCAAAATCTGTAT	5338	Db	6359	CCGACGCGTTCCGTGGCAAGCAAAAGTTCAAAATCACCAACTGGTCCACTACAACAA	6418
QY	4998	-----	4997	QY	4998	-----	4997
Db	5339	GTACATTTGCAAGATAAATAATATATCATCATGAACAATAAACTGTCTGTACATAA	5398	Db	6419	AGCTCTCATCAACCGTGGCTCCCTCACTTTCTGGCTGGATGATGGGGCGATTAGGCCCTG	6478
QY	4998	-----	4997	QY	4998	-----	5008
Db	5399	ACAGTAATAACAGGGGTGTATGAGCCATAATCAACGGGAAACGTTCTGTCAAGGCGC	5458	Db	6479	GTATGAGTCAGCAACACCTTCTTTCAGGAGGAGACCTCAGCGCCCCCCCCCTGCAGG	6538
QY	4998	-----	4997	QY	5009	CTTTCGGCGGGGTGCTCAACGGCTCAACCTACTACTGGGCTGCTTCTTAATGCAGGA	5068
Db	5459	GATTAAATCCAAATGATGCTGATTTATATGAGGTATTAATGGCTCGCGATATGCTG	5518	Db	6539	TCACACGGCGGGTCTCAACGGCTCAACCTACTACTGGGCTGCTTCTTAATGCAGGA	6598
QY	4998	-----	4997	QY	5069	GTGCGATAGGAGAGAGCGTCCAGTATCTATGATTCGAAATGGAATGGTATACCCGC	5128
Db	5519	GSCAATCAGGTGCGACAATCTATCGATTGTATGGAGCCGATGCGCCAGAGTTGTTTC	5578	Db	6599	GTGCGATAGGAGAGAGCGTCCAGTATCTATGATTCGAAATGGAATGGTATACCCGC	6658
QY	4998	-----	4997	QY	5129	ATTCTTCAGTGTCTTGGAGTCTCCTTATCAGATTTATGCCCAACTAAAGCAACCGGAGGAG	5188
Db	5579	TGAAACATGGCAAGGTAGCGTTGCGCAATGATGTTACAGATGAGATGGTCAGACTAACT	5638	Db	6659	ATTCTTCAGTGTCTTGGAGTCTCCTTATCAGATTTATGCCCAACTAAAGCAACCGGAGGAG	6718
QY	4998	-----	4997	QY	5189	AGATTTTCATGTTAAATTTCTCTGACTTTTGGTCATCAGTACGACTCGAACTGTGACACTAT	5248
Db	5639	GGCTGACGGAATTTATGCTCTTCCGACCATCAAGCAITTTTATCCGTACTCCTCATGATG	5698	Db	6719	AGATTTTCATGTTAAATTTCTCTGACTTTTGGTCATCAGTACGACTCGAACTGTGACACTAT	6778
QY	4998	-----	4997	QY	5249	CTCGGTTATGACAGCAGAAATGTCTTCTTGGAGACAGTAAATGAAGTCCCACTTAAATAA	5308
Db	5699	CATGGTTACTCACCCTCGCATCCCGGGAACACAGCATTCAGGATTTAGAGAAATATC	5758	Db	6779	CTCGGTTATGACAGCAGAAATGTCTTCTTGGAGACAGTAAATGAAGTCCCACTTAAATAA	6838
QY	4998	-----	4997	QY	5309	GAAATCTTGTGTTATCAGGNAACAACTTCTTGTGTTTTCGAACTTTTTCGGTGCCTTGAAC	5368
Db	5759	CTGATTCAGGTGAAATATTTGTATGCTGCGCTGGCAGTGTCTTCCGCGCGGTTGCAATCGA	5818	Db	6839	GAAATCTTGTGTTATCAGGNAACAACTTCTTGTGTTTTCGAACTTTTTCGGTGCCTTGAAC	6898
QY	4998	-----	4997	QY	5369	AAAACTGAGTGGATATGTCCGGTAGGAATGGAGGGGCAAACTCTTACCTTCTGGACC	5428
Db	5819	TTCTGTTTGTAAATGTCCTTTTAAACAGCGATCGCGTATTTCTGCTCGCTCAGCGCAAT	5878	Db	6899	AAAACTGAGTGGATATGTCCGGTAGGAATGGAGGGGCAAACTCTTACCTTCTGGACC	6958
QY	4998	-----	4997	QY	5429	TTCAAGAGGTATGTAGGGTTTGTAGTACTGATGCAAACTTTCAGTGAACAAGTCTTATT	5488
Db	5879	CACGAATGAAACGCTTTGTTGATGCGAGTATTTTGTATGACGAGCGTAAATCGCTGGC	5938	Db	6959	TTCAAGAGGTATGTAGGGTTTGTAGTACTGATGCAAACTTTCAGTGAACAAGTCTTATT	7018
QY	4998	-----	4997	QY	5489	TCGTTCAAAACCAATTCGAATCCAGAGAAATCAAAAGTTGTTGTCTACTATTGATTCACAGC	5548

Db 7019 TCGTTCAAAACCATCCGAATCCAGAGAAATCAAAAGTTGTTGTCTACTATTGATCCCAAGC 7078
Qy 5549 CAGTGGCGTCTTTGAAACCTGCAATAGTGTGCTGCTGTTTTCAGGTCACTCTTTGTATGAAT 5608
Db 7079 CAGTGGCGTCTTTGAAACCTGCAATAGTGTGCTGCTGTTTTCAGGTCACTCTTTGTATGAAT 7138
Qy 5609 AAATCTAGTCTTTGATCTTAAATAATCTTTGACGAGCCCAAGGCGATPAATAATACCAAACTAA 5668
Db 7139 AAATCTAGTCTTTGATCTTAAATAATCTTTGACGAGCCCAAGGCGATPAATAATACCAAACTAA 7198
Qy 5669 AACTCTTTTAAACGTTTAAAGGCAAGTATGCTGCTGATTAATAACCCCAATCAGCT 5728
Db 7199 AACTCTTTTAAACGTTTAAAGGCAAGTATGCTGCTGATTAATAACCCCAATCAGCT 7258
Qy 5729 CGTAGTCTGATCTTCATCAACTTGAAGGGCCACTATCTTGTTTTAAAGAAATTTTCGGGAGA 5788
Db 7259 CGTAGTCTGATCTTCATCAACTTGAAGGGCCACTATCTTGTTTTAAAGAAATTTTCGGGAGA 7318
Qy 5789 TGGGATATCGAGAAAAAGGTACGCTGATTTTAAAGGTGAATTTTATCTCAAGATCTCTGC 5848
Db 7319 TGGGATATCGAGAAAAAGGTACGCTGATTTTAAAGGTGAATTTTATCTCAAGATCTCTGC 7378
Qy 5849 CTCGGCGCTTTCCGGTGATGACGGTGAAAAACCTCTGACACATGACGCTCCCGGAGACGGTC 5908
Db 7379 CTCGGCGCTTTCCGGTGATGACGGTGAAAAACCTCTGACACATGACGCTCCCGGAGACGGTC 7438
Qy 5909 ACAGCTTCTGTAAAGCGGATGCGGGAGCAGACAAGCCCGTCAAGGGCGGTCAAGCGGT 5968
Db 7439 ACAGCTTCTGTAAAGCGGATGCGGGAGCAGACAAGCCCGTCAAGGGCGGTCAAGCGGT 7498
Qy 5969 GTTGGCGGTGTGCGGGGCGCAGCATGACCCAGTCAAGTACGATAGCGGAGTGTATCT 6028
Db 7499 GTTGGCGGTGTGCGGGGCGCAGCATGACCCAGTCAAGTACGATAGCGGAGTGTATCT 7558
Qy 6029 GGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGAAAA 6088
Db 7559 GGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGAAAA 7618
Qy 6089 TACCGCAGATGCGTAAAGGAAAAATACCGCATCAGCGCTCTTCCGCTTCTCGCTCA 6148
Db 7619 TACCGCAGATGCGTAAAGGAAAAATACCGCATCAGCGCTCTTCCGCTTCTCGCTCA 7678
Qy 6149 CTGACTCGCTCGCTCGCTCGCTCGCTGCGGAGCGGTATCAGCTCACCTCAAAGCGG 6208
Db 7679 CTGACTCGCTCGCTCGCTCGCTCGCTGCGGAGCGGTATCAGCTCACCTCAAAGCGG 7738
Qy 6209 TAATACGGTTATCCACAGATCAGGGGATAACGAGGAAAGAAATCATGTGAGCAAAAGGCC 6268
Db 7739 TAATACGGTTATCCACAGATCAGGGGATAACGAGGAAAGAAATCATGTGAGCAAAAGGCC 7798
Qy 6269 AGCAAAAGCCAGGAACCGTAAAGGCGCGTTCGCTGCGGTTTTTCCATAGGCTCCGCC 6328
Db 7799 AGCAAAAGCCAGGAACCGTAAAGGCGCGTTCGCTGCGGTTTTTCCATAGGCTCCGCC 7858
Qy 6329 CCCCTGACGAGCATCACAATAATCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGGAC 6388
Db 7859 CCCCTGACGAGCATCACAATAATCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGGAC 7918
Qy 6389 TATAAGATACGAGGTTTCCCGCTGGAAGTCCCTCGTGGCTCTCTCTGTTCCGACCC 6448
Db 7919 TATAAGATACGAGGTTTCCCGCTGGAAGTCCCTCGTGGCTCTCTCTGTTCCGACCC 7978
Qy 6449 TGCCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCGTTTCTCAT 6508
Db 7979 TGCCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCGTTTCTCAT 8038
Qy 6509 GCTACGCTGTAAGGTATCTCAGTTCGGTGTAGGTGCTGCTCCAGTGGGCTGTGTGC 6568
Db 8039 GCTACGCTGTAAGGTATCTCAGTTCGGTGTAGGTGCTGCTCCAGTGGGCTGTGTGC 8098
Qy 6569 AGGAACCCCGCTTACGAGCCGACCGCTGCGCTTATCCGGTAACTATCTGAGTCCA 6628

Db 8099 AGAAACCCCGCTTCAGCCCGACCGCTGCGCCTTATTCGGGTAACTATCGTCTTGAGTCCA 8158
Qy 6629 ACCCGTAAAGACACAGCTTATCGCCACTGCGCAGCAGCCACTGCTAACAGGATTAGCAGAG 6688
Db 8159 ACCCGTAAAGACACAGCTTATCGCCACTGCGCAGCAGCCACTGCTAACAGGATTAGCAGAG 8218
Qy 6689 CGAGGTATGAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTAACTACGCTACACTA 6748
Db 8219 CGAGGTATGAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTAACTACGCTACACTA 8278
Qy 6749 GAAGGACAGTATTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTG 6808
Db 8279 GAAGGACAGTATTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTG 8338
Qy 6809 GTAGTCTTGTATCCCGCAAAACAAAACACCGCTGAGTACGGGTGTTTTTTTGTGCAAGC 6868
Db 8339 GTAGTCTTGTATCCCGCAAAACAAAACACCGCTGAGTACGGGTGTTTTTTTGTGCAAGC 8398
Qy 6869 AGCAGATTACGCGCAGAAAAAGGATCTCAAGAGATCCTTTGTATCTTTCTACGGGT 6928
Db 8399 AGCAGATTACGCGCAGAAAAAGGATCTCAAGAGATCCTTTGTATCTTTCTACGGGT 8458
Qy 6929 CTGACGCTCAGTGGAAACGAAAACTCAGTTAAAGGGATTTTGTCTATGAGATTATCAAAAA 6988
Db 8459 CTGACGCTCAGTGGAAACGAAAACTCAGTTAAAGGGATTTTGTCTATGAGATTATCAAAAA 8518
Qy 6989 GGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATAT 7048
Db 8519 GGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATAT 8578
Qy 7049 ATCAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGGACCTATCTCAGCGA 7108
Db 8579 ATCAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGGACCTATCTCAGCGA 8638
Qy 7109 TCTGCTTATTTGCTTCAATCATAGTTTGCCTGACTCCCGCTGCTGTAGATAACTACGATAC 7168
Db 8639 TCTGCTTATTTGCTTCAATCATAGTTTGCCTGACTCCCGCTGCTGTAGATAACTACGATAC 8698
Qy 7169 GGGAGGCTTACCATCTGCGCCCGAGTGTGCAATGATACCGGAGACCCACCGTCAACCGG 7228
Db 8699 GGGAGGCTTACCATCTGCGCCCGAGTGTGCAATGATACCGGAGACCCACCGTCAACCGG 8758
Qy 7229 CTCAGATTATCAGCAATAAAACAGCAGCGGAAAGGGCCGAGCGCAAGAGTGTCTTG 7288
Db 8759 CTCAGATTATCAGCAATAAAACAGCAGCGGAAAGGGCCGAGCGCAAGAGTGTCTTG 8818
Qy 7289 CAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTCCGGGAAGCTAGAGTAAGTAT 7348
Db 8819 CAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTCCGGGAAGCTAGAGTAAGTAT 8878
Qy 7349 CGCAGTTAATAGTTTGCACAACGTTTGCATTTGCTGTCAGGCAATCGTGTGTCAACGCT 7408
Db 8879 CGCAGTTAATAGTTTGCACAACGTTTGCATTTGCTGTCAGGCAATCGTGTGTCAACGCT 8938
Qy 7409 GTGCTTTGGTATGGCTTCAATTCAGTCCGTTCCCAAGCATCAAGCGCAGTTACATGAT 7468
Db 8939 GTGCTTTGGTATGGCTTCAATTCAGTCCGTTCCCAAGCATCAAGCGCAGTTACATGAT 8998
Qy 7469 CCCCCATGTTGCAAAAAAGGGTTAGTCTCCTTCGGTCTCCGATCGTTGTGCAAGATA 7528
Db 8999 CCCCCATGTTGCAAAAAAGGGTTAGTCTCCTTCGGTCTCCGATCGTTGTGCAAGATA 9058
Qy 7529 AGTTGCGCGCAGTGTATCACTCATGGTTATGGCAGCATGCAATAATTCCTTTACTGTCA 7588
Db 9059 AGTTGCGCGCAGTGTATCACTCATGGTTATGGCAGCATGCAATAATTCCTTTACTGTCA 9118
Qy 7589 TGCCATCCGTAAGATCTTCTGTCAGTGGTGAAGTCAACCAAGTCAATTCCTGAGAAT 7648
Db 9119 TGCCATCCGTAAGATCTTCTGTCAGTGGTGAAGTCAACCAAGTCAATTCCTGAGAAT 9178
Qy 7649 AGTGTATGCGCGCAGCAGTGTCTTTCGCCCGCGTCAACACGGGATAATACCGCGCCAC 7708
Db 9179 AGTGTATGCGCGCAGCAGTGTCTTTCGCCCGCGTCAACACGGGATAATACCGCGCCAC 9238

QY 7709 ATAGCAGAACTTTAAAGTGTCTCATCTTGAAGAAACGTTCTTCGGGGCGGAAAACTCTCAA 7768
 Db 9239 ATAGCAGAACTTTAAAGTGTCTCATCTTGAAGAAACGTTCTTCGGGGCGGAAAACTCTCAA 9298
 QY 7769 GGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCATCTGTGACCCCAACTGATCTT 7828
 Db 9299 GGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCATCTGTGACCCCAACTGATCTT 9358
 QY 7829 CAGCATCTTTTACTTTCACACGCGTTTCTGGGTGAGCAAAACAGGAAGCAAAATGCCG 7888
 Db 9359 CAGCATCTTTTACTTTCACACGCGTTTCTGGGTGAGCAAAACAGGAAGCAAAATGCCG 9418
 QY 7889 CAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATPACTCATCTCTTCTTTTCAAT 7948
 Db 9419 CAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATPACTCATCTCTTCTTTTCAAT 9478
 QY 7949 ATTATTGAAGCAATTATCAGGGTTATTGTTCTCATGACGGATACATATTTGATGATTTT 8008
 Db 9479 ATTATTGAAGCAATTATCAGGGTTATTGTTCTCATGACGGATACATATTTGATGATTTT 9538
 QY 8009 AGAAAAATAAACAATAGGGGTTCCGGCGACATTTCCCGAAAAAGTGCCACCTGACGTCT 8068
 Db 9539 AGAAAAATAAACAATAGGGGTTCCGGCGACATTTCCCGAAAAAGTGCCACCTGACGTCT 9598
 QY 8069 AAGAAACCATTTATTCATGATGATTAACCTATTAATAATAGCGGTATCAGAGGCCCTTTC 8128
 Db 9599 AAGAAACCATTTATTCATGATGATTAACCTATTAATAATAGCGGTATCAGAGGCCCTTTC 9658
 QY 8129 GTCTTCAAGAAATTAATCTCATGTTTGACAGCTTATCATGATAAGCTGACTCATGTTGG 8188
 Db 9659 GTCTTCAAGAAATTAATCTCATGTTTGACAGCTTATCATGATAAGCTGACTCATGTTGG 9718
 QY 8189 TATTGTGAAATAGACGAGATCGGGAACACTGAAAAATAACAGTTATTATTCG 8241
 Db 9719 TATTGTGAAATAGACGAGATCGGGAACACTGAAAAATAACAGTTATTATTCG 9771

RESULT 5
 US-09-883-848A-34
 ; Sequence 34, Application US/09883848A
 ; Publication No. US20030022819A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ling, L.
 ; APPLICANT: Sanicola-Nadel, M.
 ; TITLE OF INVENTION: ANGIOGENESIS-MODULATING COMPOSITIONS AND USES
 ; FILE REFERENCE: CIBT-P01-119
 ; CURRENT APPLICATION NUMBER: US/09/883,848A
 ; CURRENT FILING DATE: 2001-06-18
 ; PRIOR APPLICATION NUMBER: 60/211,919
 ; PRIOR FILING DATE: 2000-06-16
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 34
 ; LENGTH: 10462
 ; TYPE: DNA
 ; ORGANISM: Plasmid pub116
 US-09-883-848A-34

Query Match 69.3%; Score 5709.8; DB 10; Length 10462;
 Best Local Similarity 77.4%; Pred. No. 0;
 Matches 8096; Conservative 0; Mismatches 142; Indels 2221; Gaps 8;

QY 2 GATCTAAACATCAAGACGAAAGGTTGAATGAAACCTTTTTCGCCATCCGACATCCACAGG 61
 Db 1 GATCTAAACATCAAGACGAAAGGTTGAATGAAACCTTTTTCGCCATCCGACATCCACAGG 60
 QY 62 TCCATTCACACATAAGTGCACAAACGCAACAGGAGGGGATACACTAGCAGCAGACCGTT 121
 Db 61 TCCATTCACACATAAGTGCACAAACGCAACAGGAGGGGATACACTAGCAGCAGACCGTT 120
 QY 122 GCAACGCGAGGACCTTCCACTCTCTTCTCTCAACACCTTTTTCGATCGAAAAACCA 181

Db 121 GCAACGCGAGGACCTTCCACTCTCTCTTCTCTCAACCCACTTTTGGCATCGAAAAACCA 180
 QY 182 GCCCAGTTATTGGGCTTGATTGGAGCTCGCTCATTTCCAAATTCCTTCTATTAGGCTACTAA 241
 Db 181 GCCCAGTTATTGGGCTTGATTGGAGCTCGCTCATTTCCAAATTCCTTCTATTAGGCTACTAA 240
 QY 242 CACCATGATCTTATTAGCTGTCTATCTGTGCCCCCTGGCGAGGTTTCATGTTTGTAT 301
 Db 241 CACCATGATCTTATTAGCTGTCTATCTGTGCCCCCTGGCGAGGTTTCATGTTTGTAT 300
 QY 302 TTCCGAATGCAACAGCTCCGCATTACACCCGAACATCACTCCAGATGAGGGCTTCTGA 361
 Db 301 TTCCGAATGCAACAGCTCCGCATTACACCCGAACATCACTCCAGATGAGGGCTTCTGA 360
 QY 362 GTGTGGGTCAAAATAGTTTCAATTTCCCAATAGGCCCAAACTGACAGTTTAAACGCTG 421
 Db 361 GTGTGGGTCAAAATAGTTTCAATTTCCCAATAGGCCCAAACTGACAGTTTAAACGCTG 420
 QY 422 TCTTGGAACTTAATATGACAAAGCGTGATCTCATCAAGATGAACCTAAAGTTTGGTTCGT 481
 Db 421 TCTTGGAACTTAATATGACAAAGCGTGATCTCATCAAGATGAACCTAAAGTTTGGTTCGT 480
 QY 482 TGAATATGCTAACGGCCAGTTGGTCAAAAAGAAACTTCCAAAAGTCGCATACCGTTTGTGTC 541
 Db 481 TGAATATGCTAACGGCCAGTTGGTCAAAAAGAAACTTCCAAAAGTCGCATACCGTTTGTGTC 540
 QY 542 TTGTTTGTGATTTGATTTGAGCAATGCTCAAAAATTAATCTCATTAAATGCTTAGCGAGTCTC 601
 Db 541 TTGTTTGTGATTTGATTTGAGCAATGCTCAAAAATTAATCTCATTAAATGCTTAGCGAGTCTC 600
 QY 602 TCTATCGCTTCTGAAACCCCGGTGCACTGTGCCGAAAGCAAAATGGGAAACACACCCGCTT 661
 Db 601 TCTATCGCTTCTGAAACCCCGGTGCACTGTGCCGAAAGCAAAATGGGAAACACACCCGCTT 660
 QY 662 TTTGGATGATTTATGCAATTTGCTCCACATTTGATGCTTCCAGATTTCTGGTGGGATACTG 721
 Db 661 TTTGGATGATTTATGCAATTTGCTCCACATTTGATGCTTCCAGATTTCTGGTGGGATACTG 720
 QY 722 CTGATAGCTTAACGTTTCAATGATCAAAATTTAACTTCTTAAACCCCTACTTACAGCAATA 781
 Db 721 CTGATAGCTTAACGTTTCAATGATCAAAATTTAACTTCTTAAACCCCTACTTACAGCAATA 780
 QY 782 TATAAACAGAGGAGCTGCCCTGTCTTAAACCTTTTATTCATCATTTAGCTTA 841
 Db 781 TATAAACAGAGGAGCTGCCCTGTCTTAAACCTTTTATTCATCATTTAGCTTA 840
 QY 842 CTTTCATTAATTCGAGCTGTTCCCAATTCACAGCTTTTGTATTTAAACGACTTTTAAACGAC 901
 Db 841 CTTTCATTAATTCGAGCTGTTCCCAATTCACAGCTTTTGTATTTAAACGACTTTTAAACGAC 900
 QY 902 AACTTTGAGAAGATCAAAAAACAACCTTAATTTTCGAAGGATCAAAACGATGAGATTTCTT 961
 Db 901 AACTTTGAGAAGATCAAAAAACAACCTTAATTTTCGAAGGATCAAAACGATGAGATTTCTT 960
 QY 962 CAATTTTACTGCACTTTTATTCGAGCATCTCCGCAATAGCTGCTCCAGTCAACACTA 1021
 Db 961 CAATTTTACTGCACTTTTATTCGAGCATCTCCGCAATAGCTGCTCCAGTCAACACTA 1020
 QY 1022 CAACAGAGATGAAACGGCACAATAATTCGGGCTGAAGCTGTCTATCGGTTTACTCAGATTTAG 1081
 Db 1021 CAACAGAGATGAAACGGCACAATAATTCGGGCTGAAGCTGTCTATCGGTTTACTCAGATTTAG 1080
 QY 1082 AAGGGATTTGATGTTGCTGTTTGGCCATTTTCCAGCAGCAAAATACCGGTTATTGT 1141
 Db 1081 AAGGGATTTGATGTTGCTGTTTGGCCATTTTCCAGCAGCAAAATACCGGTTATTGT 1140
 QY 1142 TTATAAATACTACTATTGCGAGCATTTGCTTAAAGAAAGAGGGGTATCTCTCGAGAAA 1201
 Db 1141 TTATAAATACTACTATTGCGAGCATTTGCTTAAAGAAAGAGGGGTATCTCTCGAGAAA 1200
 QY 1202 GAGAGG----- 1207
 Db 1201 GATGCGGACCGGCGAGGGGTTTCGGGAGAGAGAGGACCCCAAAAGCTGACCCCTTTAG 1260

QY 1208 ----- 1207
Db 1261 CCTACAGCAGTTATCCCAATGTGGCCGAGAGACCCCTAGGCGCCAGCGAAGGTATG 1320
QY 1208 ----- 1207
Db 1321 AAGGGAAGATCTCCAGAACTCCGAGCGATTAAGGAACTCACCCCAATTAACAACCCCG 1380
QY 1208 ----- 1207
Db 1381 ACATCATATTAAGGATGAAGAAACACCGAGCGGACAGGCTGATGACTCAGAGGTGTA 1440
QY 1208 ----- 1207
Db 1441 AGGACAAAGTTGAACGCTTTGGCCATCTCGGTGATGAACCACTGGCCAGGAGTGAACCTGC 1500
QY 1208 ----- 1207
Db 1501 GGGTGACCGAGGGCTGGGACGAAGATGCCACCACTCAGAGGAGTCTCTGCACTACGAGG 1560
QY 1208 ----- 1207
Db 1561 GCCGCGAGTGGACATCACCAGCTGTGACCGGACCGCAGCAAGTACGGCATGTGGCCC 1620
QY 1208 ----- 1207
Db 1621 GCCTGGCGTGGAGCCGGCTTCGACTGGGTGTACTACGAGTCCAGGSCACATATCCACT 1680
QY 1208 ----- 1207
Db 1681 GCTCGGTGAAGACAGAGAACTCGGTGGCGGCAAAATCGGAGGGGTGCAGAAAACCTACA 1740
QY 1208 ----- 1207
Db 1741 CATGCCACCGTGGCCAGCACCTGAACTCTCTGGGGGGACCGTCAGTCTTCCTCTCCCCC 1800
QY 1208 ----- 1207
Db 1801 CAAAACCAAGACACCCCTCATGATCTCCCGGACCCCTGAGGTACATGCGTGTGGTGG 1860
QY 1232 CCCAGGCCAGAGAAACATGTATCATGGCCCCCGG----- 1269
Db 1861 ACGTGAGCCACGAAGACCTGAGGTCAAGTTCAACTGTGACGTGACGGCGCGGTGAGGTGC 1920
QY 1270 ----- 1294
Db 1921 ATAAATGCCAAGACAAAGCCGCGGAGGAGCAGTACCAGACAGCTACCGTGTGTGTCAGCG 1980
QY 1295 GTGTCAACCGCCAGCA----- 1310
Db 1981 TCCTCACCGTCTGCAACCCAGCACTGGCTGAATGGCAAGGATACAAGTGAAGGTCTCCA 2040
QY 1311 ----- 1310
Db 2041 ACAAAAGCCCTCCAGCCCCATCGAGAAACCATCTCCAAAGCCAAAGGGCAGCCCGAG 2100
QY 1311 ----- 1310
Db 2101 AACCAAGGTGTACACCCCTGCCCCCATCCCGGATGAGCTGACCAAGAACCAAGGTCAAGC 2160
QY 1311 ----- 1366
Db 2161 TGACCTGCTGTCAAAAGCCTTCTATCCAGCEACATCGCCGTGAGTGGGAGAGCAATG 2220
QY 1367 TCCACCCCATGGCCATCGA----- 1385
Db 2221 GGCAGCCGAGAGAACTACAAGACCAAGCCCTCCCGTGTGGACTCCGAGCGCTCCTTCT 2280
QY 1386 ----- 1385
Db 2281 TCCTCTACAGCAAGCTCACCGTGGAGAGAGAGCGTGGCAGCAGGGGAACGTTCTCTCAT 2340

QY 1386 ----- 1423
Db 2341 GCTCCGTGATGATGAGGCTCTGCACACCACTACACGAGAGAGGCTCTCCCTGTCTC 2400
QY 1424 CBTAGAATTCCTTAGGCGCGCGCAATTAATTCGCTTTAGACATGTTCTTCCTCAGTT 1483
Db 2401 CCGGAAATGAGTGCAGCGCGCGCAATTAATTCGCTTTAGACATGTTCTTCCTCAGTT 2460
QY 1484 CAAATGGGCACTTACGAGAGACCGGCTCTCTAGATTCTAATCAAGAGGATGTAGAA 1543
Db 2461 CAAATGGGCACTTACGAGAGACCGGCTCTCTAGATTCTAATCAAGAGGATGTAGAA 2520
QY 1544 TGCCATTTGCCTGAGAGATGCGAGCTTCATTTTGTACTTTTTTATTTGTAACTATAT 1603
Db 2521 TGCCATTTGCCTGAGAGATGCGAGCTTCATTTTGTACTTTTTTATTTGTAACTATAT 2580
QY 1604 AGTATAGGATTTTTTTGTCAATTTTCTCTCGTACGAGCTTCCTCTCATCAGCT 1663
Db 2581 AGTATAGGATTTTTTTGTCAATTTTCTCTCGTACGAGCTTCCTCTCATCAGCT 2640
QY 1664 ATCTCCAGCTGATGAATATCTTGTGTAGGGGTTTGGGAAATCATTCGAGTTTGTATGT 1723
Db 2641 ATCTCCAGCTGATGAATATCTTGTGTAGGGGTTTGGGAAATCATTCGAGTTTGTATGT 2700
QY 1724 TTTTCTGTGATTTTCCACTCTCTTCAGAGTACAGAGATTAAGTGAGAGTTCGTTTG 1783
Db 2701 TTTTCTGTGATTTTCCACTCTCTTCAGAGTACAGAGATTAAGTGAGAGTTCGTTTG 2760
QY 1784 TGCAAGCTTATCGATAAGCTTTAATCGGTAGTTTATCACAGTTAAATTTGCTAAACAGT 1843
Db 2761 TGCAAGCTTATCGATAAGCTTTAATCGGTAGTTTATCACAGTTAAATTTGCTAAACAGT 2820
QY 1844 CAGGACCCGTGTATGAAATCTAAACAATGCGCTCATCGTTCATCTCGGCAACCGTCACTCG 1903
Db 2821 CAGGACCCGTGTATGAAATCTAAACAATGCGCTCATCGTTCATCTCGGCAACCGTCACTCG 2880
QY 1904 GATGCTGTAGGATAGCTTGGTTATGCGGTACTGCGGGCTCTTGGGGATATCCTC 1963
Db 2881 GATGCTGTAGGATAGCTTGGTTATGCGGTACTGCGGGCTCTTGGGGATATCCTC 2940
QY 1964 CATTCGACAGCATCGCCAGTCACTATGCGGTCTGCTAGCGCTATATGCGTTGATGCA 2023
Db 2941 CATTCGACAGCATCGCCAGTCACTATGCGGTCTGCTAGCGCTATATGCGTTGATGCA 3000
QY 2024 TTTCTATGCGCACCGTTCGAGGACATGTTCGACCGCTTTGGCGCGCGCCAGTCCCTG 2083
Db 3001 TTTCTATGCGCACCGTTCGAGGACATGTTCGACCGCTTTGGCGCGCGCCAGTCCCTG 3060
QY 2084 CTGCTTTCGCTACTTGGAGCCACTATCGACTACGCGATCATGGCGACCAACACCGCTCCTG 2143
Db 3061 CTGCTTTCGCTACTTGGAGCCACTATCGACTACGCGATCATGGCGACCAACACCGCTCCTG 3120
QY 2144 TGGATCTATCGAATCTAAATGTAAATTAATCTCTAAATTAATTAATTAATTAATTAATTAAT 2203
Db 3121 TGGATCTATCGAATCTAAATGTAAATTAATCTCTAAATTAATTAATTAATTAATTAATTAAT 3180
QY 2204 CTCTATACGAACTTAAACGATTCGAGTTCGAGCTTAGACCTTCAACAGCAGCCAGATC 2263
Db 3181 CTCTATACGAACTTAAACGATTCGAGTTCGAGTTCGAGCTTAGACCTTCAACAGCAGCCAGATC 3240
QY 2264 CATCACTGTCTGGCCAAATATGTTTTCAGTCCCTCAGGAGTTACGTTGTGAAGTGTGA 2323
Db 3241 CATCACTGTCTGGCCAAATATGTTTTCAGTCCCTCAGGAGTTACGTTGTGAAGTGTGA 3300
QY 2324 CTTCTGGAAGTTGCGAGTGTAACTCGCTGTATTAACGGGCATATCCGTACGTTGGCAA 2383
Db 3301 CTTCTGGAAGTTGCGAGTGTAACTCGCTGTATTAACGGGCATATCCGTACGTTGGCAA 3360
QY 2384 AGTGTGTTGTTCCGAGGAGTAACTCCACAACCTCTCTGGAGAGTAGGACCAACAA 2443
Db 3361 AGTGTGTTGTTCCGAGGAGTAACTCCACAACCTCTCTGGAGAGTAGGACCAACAA 3420
QY 2444 CACAGATCCAGCGTGTGTACTTGATCAACATAAGAGAAAGCATTTCTCGATTTTGCAGGAT 2503

Db	7799	 TGTTTTGAGGTCATCTTTGTGTAAATAAATCTAGTCTTTTGATCTAAATATATCTTGACGAG	7858
Qy	5643	CCAGGCGGATAAATACCCAAATCTAAACCTTTTAAACGTTTAAAGGACAGTAGTATGC	5702
Db	7859	CCAGGCGGATAAATACCCAAATCTAAACCTTTTAAACGTTTAAAGGACAGTAGTATGC	7918
Qy	5703	TGCGTGTATTAACCCCAATCAGCTCGTAGTCTGATCTCTCATCAACTTTGAGGGGCACTA	5762
Db	7919	TGCGTGTATTAACCCCAATCAGCTCGTAGTCTGATCTCTCATCAACTTTGAGGGGCACTA	7978
Qy	5763	TCCTGTTTTAGAGNAATTTCCGAGATGCGATATCGAGAAAGGTACGCTGATTTTAA	5822
Db	7979	TCCTGTTTTAGAGNAATTTCCGAGATGCGATATCGAGAAAGGTACGCTGATTTTAA	8038
Qy	5823	CGTGAAATTTATCTCAAGATCTCTGCCCTCCGGGTTTCGGTGATGACGGTGAACCTCT	5882
Db	8039	CGTGAAATTTATCTCAAGATCTCTGCCCTCCGGGTTTCGGTGATGACGGTGAACCTCT	8098
Qy	5883	GACATGCGAGCTCCGAGAGCGGTACAGCTTGTCTGTAAGCGGATGCCGGAGCAGAC	5942
Db	8099	GACATGCGAGCTCCGAGAGCGGTACAGCTTGTCTGTAAGCGGATGCCGGAGCAGAC	8158
Qy	5943	AAGCCCGTCAGGGCGGTACAGCGGTGTTCCGGGTGTCGGGCGCAGCCATCACCAGT	6002
Db	8159	AAGCCCGTCAGGGCGGTACAGCGGTGTTCCGGGTGTCGGGCGCAGCCATCACCAGT	8218
Qy	6003	CAGTAGCGATAGCGGAGTGATATCGGCTTAATATCGGCGATCAGAGCAGATTGTACT	6062
Db	8219	CAGTAGCGATAGCGGAGTGATATCGGCTTAATATCGGCGATCAGAGCAGATTGTACT	8278
Qy	6063	GAGAGTCGCATATGCGGTGCGAATACCGCACAGATCGTAAGAGAGAAATACCGCAT	6122
Db	8279	GAGAGTCGCATATGCGGTGCGAATACCGCACAGATCGTAAGAGAGAAATACCGCAT	8338
Qy	6123	CAGGCGCTCTTCCGCTTCTCGTCACTGACTCGCTCGCTCGGTGTTCCGGCTCGGGG	6182
Db	8339	CAGGCGCTCTTCCGCTTCTCGTCACTGACTCGCTCGGTGTTCCGGCTCGGGG	8398
Qy	6183	AGCGGTATCAGTCACTCAAAGGCGGTAAATACGGTTATCCACAGATCAGGGGATAACG	6242
Db	8399	AGCGGTATCAGTCACTCAAAGGCGGTAAATACGGTTATCCACAGATCAGGGGATAACG	8458
Qy	6243	AGGAAAGAACATGTAGCAAAAGGCCAGCAAAAGGCCAGAAACGTAAGAGCCCGCTT	6302
Db	8459	AGGAAAGAACATGTAGCAAAAGGCCAGCAAAAGGCCAGAAACGTAAGAGCCCGCTT	8518
Qy	6303	GTCGGCGTTTTTCATAGGTCCGCCCTCAGCAGCATCAAAATTCGACGCTCAAG	6362
Db	8519	GTCGGCGTTTTTCATAGGTCCGCCCTCAGCAGCATCAAAATTCGACGCTCAAG	8578
Qy	6363	TCAGAGTGGCGAAACCGCAGGACTATAAAGATACAGCGGTTTCCCGCTCGAAGCTC	6422
Db	8579	TCAGAGTGGCGAAACCGCAGGACTATAAAGATACAGCGGTTTCCCGCTCGAAGCTC	8638
Qy	6423	CCTCGTCGCTCTCCTGTTCCGACCCCTCCCGCTTACCGGATACTGTCGCGCTTTCTCC	6482
Db	8639	CCTCGTCGCTCTCCTGTTCCGACCCCTCCCGCTTACCGGATACTGTCGCGCTTTCTCC	8698
Qy	6483	TTCCGGAAGCGTGCGCTTTCTCATAGCTCAGCTGTAGTATCTCAGTTTCGGGTAGGT	6542
Db	8699	TTCCGGAAGCGTGCGCTTTCTCATAGCTCAGCTGTAGTATCTCAGTTTCGGGTAGGT	8758
Qy	6543	CGTTTCGCTCCAAGCTGGGCTGTGTGCAGAACCCCGCTTCAGCCCGACCGCTGCGCTT	6602
Db	8759	CGTTTCGCTCCAAGCTGGGCTGTGTGCAGAACCCCGCTTCAGCCCGACCGCTGCGCTT	8818
Qy	6603	ATCCGGTAACTATCGTCTTGAATCCACCGGTAAAGCAGACTTATCGGCATCTGCAGC	6662
Db	8819	ATCCGGTAACTATCGTCTTGAATCCACCGGTAAAGCAGACTTATCGGCATCTGCAGC	8878
Qy	6663	AGCCACTGGTAAACAGGATTAGCAGAGCGGATGTAGGCGGTGCTACAGATTCTTTGAA	6722

8879	AGCCACTGCTAAACAGGATTAGCAGAGCGAGGATATAGTAGCGGGTGCTACAGAGTTCITGAA	8938
6723	GTGGTGGCCCTAACTACGGCTACACTAGAGGACAGATTTTGGTATCTGCGCTCTGCTGAA	6782
8939	GTGGTGGCCCTAACTACGGCTACACTAGAGGACAGATTTTGGTATCTGCGCTCTGCTGAA	8998
6783	GCCAGTTACCTTCGGHAAAAGAGTTGGTAGCTCTTGATCGGCAAAACAACACACCGCTGG	6842
8999	GCCAGTTACCTTCGGHAAAAGAGTTGGTAGCTCTTGATCGGCAAAACAACACACCGCTGG	9058
6843	TAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCCGAAAAAAGATCTCTCAAGA	6902
9059	TAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCCGAAAAAAGATCTCTCAAGA	9118
6903	AGATCCTTTTGATCTTTTCTACGGGGTCTGACGCTAGTGGACGAAACCTCAGTTTAAAG	6962
9119	AGATCCTTTTGATCTTTTCTACGGGGTCTGACGCTAGTGGACGAAACCTCAGTTTAAAG	9178
6963	GATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAATG	7022
9179	GATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAATG	9238
7023	AAATTTTAAATCAATCTAAAGTATATATCAGTAAACTTGGTCTGA CAGTTACCAATGCTTT	7082
9239	AAATTTTAAATCAATCTAAAGTATATATCAGTAAACTTGGTCTGA CAGTTACCAATGCTTT	9298
7083	AATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCTTCATCCTAGTTGCCTGACT	7142
9299	AATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCTTCATCCTAGTTGCCTGACT	9358
7143	CCCGGTGCTGATAACTACGATACGGGAGGCTTACCATCTGGCCCCCAGTCTGCTCAAT	7202
9359	CCCGGTGCTGATAACTACGATACGGGAGGCTTACCATCTGGCCCCCAGTCTGCTCAAT	9418
7203	GATACCGCAGAGCCACGCTCACCGGCTCCAGATTTTATCAGCAATTAACACGACGACCGG	7262
9419	GATACCGCAGAGCCACGCTCACCGGCTCCAGATTTTATCAGCAATTAACACGACGACCGG	9478
7263	AAGGGCCGAGCGAGAAAGTGGTCTCCCAACTTTATCCGCTCCATCCAGTCTATTAAATG	7322
9479	AAGGGCCGAGCGAGAAAGTGGTCTCCCAACTTTATCCGCTCCATCCAGTCTATTAAATG	9538
7323	TTGCCGGGAAGCTAGAGTAGTTTGGCGAGTTAATAGTTTGGCGAAAGTGGTGGCCAT	7382
9539	TTGCCGGGAAGCTAGAGTAGTTTGGCGAGTTAATAGTTTGGCGAAAGTGGTGGCCAT	9598
7383	TGCTGCAGGCATCGTGGTGTACGCTCGTCTGTTGGTATGGCTTCATTACGCTCCGGTTC	7442
9599	TGCTGCAGGCATCGTGGTGTACGCTCGTCTGTTGGTATGGCTTCATTACGCTCCGGTTC	9658
7443	CCAACGATCAAGCGAGTTACATGATCCGCCCATGTTTGTGCAAAAAGCGGTTAGCTCCTTT	7502
9659	CCAACGATCAAGCGAGTTACATGATCCGCCCATGTTTGTGCAAAAAGCGGTTAGCTCCTTT	9718
7503	CGGTCTCCGATCGTTGTACAGAGTAGTTGGCCGCGAGTGTATCACTCATGTTATGGC	7562
9719	CGGTCTCCGATCGTTGTACAGAGTAGTTGGCCGCGAGTGTATCACTCATGTTATGGC	9778
7563	AGCACTGCATAATTCCTTTATCTGTCAATGCCATCCGTAAGATGTTTTTCTGTGATGCTGA	7622
9779	AGCACTGCATAATTCCTTTACTGTCTATGTCATGCCATCCGTAAGATGTTTTTCTGTGATGCTGA	9838
7623	GTACTCAACCAAGTCAATCTGAGAAATAGTGTATGCGCGACCGAGTGTCTTTGCCCGGC	7682
9839	GTACTCAACCAAGTCAATCTGAGAAATAGTGTATGCGCGACCGAGTGTCTTTGCCCGGC	9898
7683	GTCAACACGGGATAATACCGCGCCCATAGCAAGACTTTTAAAGTGTCTCATCATTTGGAAA	7742
9899	GTCAACACGGGATAATACCGCGCCCATAGCAAGACTTTTAAAGTGTCTCATCATTTGGAAA	9958
7743	ACGTTCTTCGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCGATGA	7802
9959	ACGTTCTTCGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCGATGA	1001

QY	7803	ACCCACTCGTGCACCAACTGATCTTTCAGCATCTTTTACTTTTCACCAAGCGTTTCTGGGTG	7862
Db	10019	ACCCACTCGTGCACCAACTGATCTTTCAGCATCTTTTACTTTTCACCAAGCGTTTCTGGGTG	10078
QY	7863	AGCAAAAAAGGAGGCAAAATGCCGCAAAAAGGAATAAGGGCGACACGGAATCTTG	7922
Db	10079	AGCAAAAAAGGAGGCAAAATGCCGCAAAAAGGAATAAGGGCGACACGGAATCTTG	10138
QY	7923	AATACTCATACTCTTCCTTTTTTCAATATATTATGAAGCACTTTATCAGGGTTATTCTCTCAT	7982
Db	10139	AATACTCATACTCTTCCTTTTTTCAATATATTATGAAGCACTTTATCAGGGTTATTCTCTCAT	10198
QY	7983	GAGCGGATACATATTGGAATGTATTTAGAAAAATAAACAATAGGGTTCCGCGCACAATT	8042
Db	10199	GAGCGGATACATATTGGAATGTATTTAGAAAAATAAACAATAGGGTTCCGCGCACAATT	10258
QY	8043	TCCCGAAAAAGTGCACACTTGAGCTCTTAAGAAACCACTTATTATCATGACATTAAACCTATAA	8102
Db	10259	TCCCGAAAAAGTGCACACTTGAGCTCTTAAGAAACCACTTATTATCATGACATTAAACCTATAA	10318
QY	8103	AAATAGCGGTATCACAGAGCCCTTTTCGTCTTCAAGAAATAAATCTCATGTTTGACAGCTT	8162
Db	10319	AAATAGCGGTATCACAGAGCCCTTTTCGTCTTCAAGAAATAAATCTCATGTTTGACAGCTT	10378
QY	8163	ATCATCGATAAGCTGACTCATGTTGTGTAATGAAATAGACGACATCGGAAACACTGAA	8222
Db	10379	ATCATCGATAAGCTGACTCATGTTGTGTAATGAAATAGACGACATCGGAAACACTGAA	10438
QY	8223	AAATAACAGTTATTATTCG	8241
Db	10439	AAATAACAGTTATTATTCG	10457
RESULT 10			
US-09-883-848A-32			
; Sequence 32, Application US/09883848A			
; Publication No. US20030022819A1			
; GENERAL INFORMATION:			
; APPLICANT: Ling, L.			
; APPLICANT: Sanicola-Nadel, M.			
; TITLE OF INVENTION: ANGIOGENESIS-MODULATING COMPOSITIONS AND USES			
; FILE REFERENCE: CIBT-P01-119			
; CURRENT APPLICATION NUMBER: US/09/883,848A			
; CURRENT FILING DATE: 2001-06-18			
; PRIOR APPLICATION NUMBER: 60/211,919			
; PRIOR FILING DATE: 2000-06-16			
; NUMBER OF SEQ ID NOS: 48			
; SOFTWARE: PatentIn ver. 2.1			
; SEQ ID NO 32			
; LENGTH: 10491			
; TYPE: DNA			
; ORGANISM: Plasmid PUB114			
US-09-883-848A-32			
Query Match 68.8%; Score 5655.8; DB 10; Length 10491;			
Best Local Similarity 77.2%; Pred. No. 0;			
Matches 8100; Conservative 0; Mismatches 137; Indels 2252; Gaps 10;			
QY	2	GATCTACATCCAAAGACGAAGGTTGAATGAACCTTTTGGCCATCCGACATCCACAGG	61
Db	1	GATCTACATCCAAAGACGAAGGTTGAATGAACCTTTTGGCCATCCGACATCCACAGG	60
QY	62	TCCATTCTCACATAAGTGCCAAAACGCAACGAGGGGATACACTAGCAGACCGTT	121
Db	61	TCCATTCTCACATAAGTGCCAAAACGCAACGAGGGGATACACTAGCAGACCGTT	120
QY	122	GCAACCGAGGACTCCACTCTCTCTCTCTCAACACCCACTTTTGGCCATCGAAAAACA	181
Db	121	GCAACCGAGGACTCCACTCTCTCTCTCTCTCAACACCCACTTTTGGCCATCGAAAAACA	180
QY	182	GCCCAGTTATTGGCGTTGATTGGAGCTCGCTCAATTCCTTATTAGGCTACTAA	241

QY 1243 GAAGAAACAT----- 1252
Db 1320 AAGGAAGATCTCCAGAACTCCGAGGATTTAAGGAACCTCACCCCAATTAACAACCCG 1379
QY 1253 -----GTAATATGCCCCCGGAGAGGATAAATTTGGCTTCCCGGTTGC 1299
Db 1380 ACATCATATTTAAGGATGAAGAAAACACCGAGCGGACAGCTGATGACTCAGAGGTGA 1439
QY 1300 A----- 1300
Db 1440 AGACAAGTTGAACGCTTTGGCCATCTCGGTGATGAACCACTGGCCAGGAGTGAACATGC 1499
QY 1301 ----- 1300
Db 1500 GGGTGACCGAGGGCTGGGACGAAGATGGCCACACTCAGAGGAGTCTCTGCACTACGAGG 1559
QY 1301 ----- 1300
Db 1560 GCCGCGCAGTGACATCACCACTGTGACCGGACCGCAGCAAGTACGGCATGTGGCCCC 1619
QY 1301 ----- 1300
Db 1620 GCCTGGCGGTGGAGGCGGCTTCGACTCGGTGTACTAGGAGTCCAAAGGCACATATCCACT 1679
QY 1301 ----- 1300
Db 1680 GCTCGGTGAACGACAGAACTCGGTGGCGGCCAAATCGGAGGCGGTGACGTGCCAGGG 1739
QY 1301 ----- 1300
Db 1740 ATTGTGTTGTAAGCCTTGCAATATGACAGTCCGAGAGTATCATCTGCTCTCATCTTC 1799
QY 1301 ----- 1300
Db 1800 CCCCAGGCCAAGGATGTCTCACATTACTCTGACTCTTAAGTCACTGCTGTGTGTGG 1859
QY 1301 ----- 1300
Db 1860 TAGACATCAGCAAGGATGATCCGAGGTCAGTTTCAGCTGGTTCTAGATGATGGAGG 1919
QY 1301 ----- 1300
Db 1920 TGCACACAGCTCAGACGCAACACCGGAAGAGCAGTTCCAAAGCAGCTTTCGCTCAGTCA 1979
QY 1301 ----- 1300
Db 1980 GTGAATCTCCCATCATGCACAGGACTGGCTCAATGGCAAGGAGTTCAATGCAAGGTCA 2039
QY 1301 ----- 1300
Db 2040 ACAGTGCACTTTCCTGCCCCCATCGAGAAAACCATCTCCAAAACCAAGGCAGACCGA 2099
QY 1301 -----CCG 1303
Db 2100 AGGCTCCAGGTGACACCATTCACCTCCCAAGGACGAGATGGCCAGGATAAAGTCA 2159
QY 1304 CCCAGCAGTGCCAGGAGAGGTTGCTGTTTGTGATGACAGTGTCCGGGGAATTCCTCGTGT 1363
Db 2160 GTCTGACCTGCATGATAACAGACTTCTTCCCTGAAGACATTACTGTGGAGTGGCAGTGA 2219
QY 1364 GCTTCCACCCATGGCCATCGAACA----- 1390
Db 2220 ATGGCAGCCGCGAGAGACTACAAGAACACTCAGCCCATCATGGACACAGATGGCTT 2279
QY 1391 ----- 1390
Db 2280 ACTTCGTCTACAGCAAGCTCAATGTGCAGAGAGCAACTGGGAGGAGGAGGAAATACTTTCA 2339
QY 1391 -----CTCAAGAAAGAAATGTCCTCTTAACT 1418
Db 2340 CCGTCTCTGTGTATAGAGGCGCTGCACAAACCACTACTGAGAGAGGCGCTCTCCCACT 2399

QY 1419 AGTGGCGTAGA-----ATTCCCTAGGGCGCGCGCAATTA 1453
Db 2400 CTCCTGTAAATGATCCCAAGTGTCCCTGGAGCCCTCTGTCTCAGCGCGCGCAATTA 2459
QY 1454 ATTCCGCTTAGACATGACATGTTCCCTCAGTTCAAGTTGGGCACCTTACGAGAACACCGTCT 1513
Db 2460 ATTCCGCTTAGACATGACATGTTCCCTCAGTTCAAGTTGGGCACCTTACGAGAACACCGTCT 2519
QY 1514 TGCTAGATTCTAATCAAGAGGATGTCAGAATGCCATTTGCCCTGAGAGATGCAAGGTTCAAT 1573
Db 2520 TGCTAGATTCTAATCAAGAGGATGTCAGAATGCCATTTGCCCTGAGAGATGCAAGGTTCAAT 2579
QY 1574 TTTTGATACCTTTTATTTGTAACCTATATAGTATAGGATTTTTTTTGTCAATTTGTTTC 1633
Db 2580 TTTTGATACCTTTTATTTGTAACCTATATAGTATAGGATTTTTTTTGTCAATTTGTTTC 2639
QY 1634 TTCTCGTACGAGTTGCTCTCTGATCAGCCCTATCTCCAGCTCATGAATATCTTGTGGTAG 1693
Db 2640 TTCTCGTACGAGTTGCTCTCTGATCAGCCCTATCTCCAGCTCATGAATATCTTGTGGTAG 2699
QY 1694 GGGTTTGGGAAATCATTTGAGTTTGAATTTTTCTTGGTATTTCCCACTCTCTTCAGA 1753
Db 2700 GGGTTTGGGAAATCATTTGAGTTTGAATTTTTCTTGGTATTTCCCACTCTCTTCAGA 2759
QY 1754 GTACAGAAATTAAGTGAGAGTTCGTTTGTGCAAGCTTATCGATAAGCTTTAATGCCGT 1813
Db 2760 GTACAGAAATTAAGTGAGAGTTCGTTTGTGCAAGCTTATCGATAAGCTTTAATGCCGT 2819
QY 1814 AGTTATCAAGTTAAATTCGCTAACGAGTACGACCGCTGATGATGAATCTAACAATGCG 1873
Db 2820 AGTTATCAAGTTAAATTCGCTAACGAGTACGACCGCTGATGATGAATCTAACAATGCG 2879
QY 1874 CTCATCGTATCTCGGCACCGTCACTCGGATGCTGTAGGCATAGGCTTGGTTATGCGG 1933
Db 2880 CTCATCGTATCTCGGCACCGTCACTCGGATGCTGTAGGCATAGGCTTGGTTATGCGG 2939
QY 1934 GTATCGCGGGCTCTTGGGATATCGTCCATTCGACAGCATCGCAGTCACTATGCG 1993
Db 2940 GTATCGCGGGCTCTTGGGATATCGTCCATTCGACAGCATCGCAGTCACTATGCG 2999
QY 1994 GTGCTGCTAGCCCTATATCGGTTGATGCAATTTCTATGCGCACCCGTTCTCGAGCAGCTG 2053
Db 3000 GTGCTGCTAGCCCTATATCGGTTGATGCAATTTCTATGCGCACCCGTTCTCGAGCAGCTG 3059
QY 2054 TCCGACCGCTTTGGCGCGCGCCAGTCTGCTGCTCGTACTTTGGAGCCACTATCGAC 2113
Db 3060 TCCGACCGCTTTGGCGCGCGCCAGTCTGCTGCTCGTACTTTGGAGCCACTATCGAC 3119
QY 2114 TAGCGATCATGGCGACACACCGCTCTGAGTATCGAATCTAATGTAAGTTAAA 2173
Db 3120 TAGCGATCATGGCGACACACCGCTCTGAGTATCGAATCTAATGTAAGTTAAA 3179
QY 2174 ATCTCTAAATTAATTAAGTCCAGTTTCTCCATAGAACTTAAACAGCATTCGGGTG 2233
Db 3180 ATCTCTAAATTAATTAAGTCCAGTTTCTCCATAGAACTTAAACAGCATTCGGGTG 3239
QY 2234 AGCATCTAGACCTTCAACAGCAGCAGATCCATCATCTGCTTGGCCAAATATGTTTCACTCC 2293
Db 3240 AGCATCTAGACCTTCAACAGCAGCAGATCCATCATCTGCTTGGCCAAATATGTTTCACTCC 3299
QY 2294 CTCAGAGTTAGCTCTTGTGAAAGTATGAATCTTGGAAAGTTGAGTGTGTTAACTCCGCT 2353
Db 3300 CTCAGAGTTAGCTCTTGTGAAAGTATGAATCTTGGAAAGTTGAGTGTGTTAACTCCGCT 3359
QY 2354 GTATTGACGGGATATTCGTAAGTTGGCAAGTGTGGTTGTTACCGGAGGAGTAACTTCC 2413
Db 3360 GTATTGACGGGATATTCGTAAGTTGGCAAGTGTGGTTGTTACCGGAGGAGTAACTTCC 3419
QY 2414 ACNACTCTCGAGAGTAGGCGACCAACACACATCCAGCGTGTGTACTTGTATCAAC 2473
Db 3420 ACNACTCTCTGAGAGTAGGCGACCAACACACATCCAGCGTGTGTACTTGTATCAAC 3479
QY 2474 ATAAGAAAGAACATTTCTCGATTTGCGAGTCAAGTGTTCAGGAGCGTACTGATTTGACAT 2533

3480 ATAAAGAGAGCAATTCGATTTGAGGATCAAGTTCAGGAGGATCAATTCGACAT 3539
2534 TTCCAAAGCCTCTCGTAGGTTGCAACCGATAGAGGTTGTAGAGTGTGCAATACACTTGGC 2593
3540 TTCCAAAGCCTCTCGTAGGTTGCAACCGATAGAGGTTGTAGAGTGTGCAATACACTTGGC 3599
2594 TACAATTTCAACCCCTTGGCAACTGCACAGCTTGGTTGTGAACAGCATCTTCAATTTCTGGC 2653
3600 TACAATTTCAACCCCTTGGCAACTGCACAGCTTGGTTGTGAACAGCATCTTCAATTTCTGGC 3659
2654 AAGCTCTCTGTCTGTATATGACAGGCAACAGAAATCACTTGGGAATCAATACCAATGTTTC 2713
3660 AAGCTCTCTGTCTGTATATGACAGGCAACAGAAATCACTTGGGAATCAATACCAATGTTTC 3719
2714 AGCTTTGAGCAGAGAGGCTTGAGGCAACGAAATCTGGATCAGCGTATTTATCAGCAATAAC 2773
3720 AGCTTTGAG-CAGAGGCTTGAGGCAACGAAATCTGGATCAGCGTATTTATCAGCAATAAC 3778
2774 TAGAACTTCAGAGGCCCCAGCAGGATGTCAATACACAGGCGTGATGTGATTTTG 2833
3779 TAGAACTTCAGAGGCCCCAGCAGGCA-TGTCAATACACAGGCGTGATGTGATTTTG 3838
2834 AACCATCATCTTGGCAGCGATGAACGAACTGGTTTCTGGACCAATATTTTGTACACTT 2893
3839 AACCATCATCTTGGCAGCAGTAACGAACTGGTTTCTGGACCAATATTTTGTACACTT 3898
2894 AGCAACAGTTTCTGTTCGTAGGCAATAGCAGTACTGCTGGGCGCTCTGCTGTAGCAC 2953
3899 AGCAACAGTTTCTGTTCGTAGGCAATAGCAGTACTGCTGGGCGCTCTGCTGTAGCAC 3958
2954 GATACACTTAGCAACCACTTGTGGGCAACGATAGATCTTCTGGGTTAAGGTTACCATC 3013
3959 GATACACTTAGCAACCACTTGTGGGCAACGATAGATCTTCTGGGTTAAGGTTACCATC 4018
3014 CTCTTTAGGTGAGATGCAAAACAAATTTCTTTGCAACGACCACTTTGGCAGGAACACC 3073
4019 CTCTTTAGGTGAGATGCAAAACAAATTTCTTTGCAACGACCACTTTGGCAGGAACACC 4078
3074 CAGCATCAGGGAAGTGAAGGAGCAATTCGGTTCCACAGGAATATAGAGGCCCAACTTT 3133
4079 CAGCATCAGGGAAGTGAAGGAGCAATTCGGTTCCACAGGAATATAGAGGCCCAACTTT 4138
3134 CTCAATAGTCTTGAACAGAGAGCAGATACACAGGCGAGTCTCAACTTGCACGT 3193
4139 CTCAATAGTCTTGAACAGAGAGCAGATACACAGGCGAGTCTCAACTTGCACGT 4198
3194 CTCCGTTAGTTCAGCTTCATGGAATTCCTGACGTTATCTATAGAGAGATCAATGGCTCT 3253
4199 CTCCGTTAGTTCAGCTTCATGGAATTCCTGACGTTATCTATAGAGAGATCAATGGCTCT 4258
3254 CTTAACGTTATCTGGCAATTCATAGTTCCTCTGGGAAAGAGCTTCTAAACAGAGTGT 3313
4259 CTTAACGTTATCTGGCAATTCATAGTTCCTCTGGGAAAGAGCTTCTAAACAGAGTGT 4318
3314 CTTCAAAGGAGCTCCATCAAACTTCGCAAGTTCCTAAAGGGCTTTGTCAACCAATTTG 3373
4319 CTTCAAAGGAGCTCCATCAAACTTCGCAAGTTCCTAAAGGGCTTTGTCAACCAATTTG 4378
3374 ACBAACATTCGCAAAATTTGGTTGACTTAATTCCTATCTGTTTCTGGATAGG 3433
4379 ACBAACATTCGCAAAATTTGGTTGACTTAATTCCTATCTGTTTCTGGATAGG 4438
3434 ACBAACATTCGCAAAATTTGGTTGACTTAATTCCTATCTGTTTCTGGATAGG 3493
4439 ACBAACATTCGCAAAATTTGGTTGACTTAATTCCTATCTGTTTCTGGATAGG 4498
3494 TTCAATACAGCTTCAGAGGAGCTCTTTTAGGTTGGATCTCTTTAGGTTGTTTCTT 3553
4499 TTCAATACAGCTTCAGAGGAGCTCTTTTAGGTTGGATCTCTTTAGGTTGTTTCTT 4558
3554 GGTGTATCTGCGTGGCAATCTCTTCTCTAGTGAACCTTTTAGGGACTTCATATCCAG 3613

4559 GGTGTATCTGCGTGGCAATCTCTTCTCTAGTGAACCTTTAGGGACTTCATATCCAG 4618
3614 GTTTCTCTCCACCTCTGTCACACGTCACCGTACTTGGACATCTAACTAATCAAAATA 3673
4619 GTTTCTCTCCACCTCTGTCACACGTCACCGTACTTGGACATCTAACTAATCAAAATA 4678
3674 AAATAAGTCAGACATTTCCAGGCTATATCTTCTCTGGATTTAGCTTCTGCAAGTTCATC 3733
4679 AAATAAGTCAGACATTTCCAGGCTATATCTTCTCTGGATTTAGCTTCTGCAAGTTCATC 4738
3734 AGCTTCTCTCTAAATTTTAGGTTTC-AACAAACTTCTGCTGCTCAAAATACCGTTGGTAT 3792
4739 AGCTTCTCTCTAAATTTTAGGTTTCACAAACAAACTTCTGCTGCTCAAAATACCGTTGGTAT 4798
3793 AAGAACCTTCTGAGCATCTCTTACGATCCCAAGGTGGCTTCCATGGCTCTTAAGAC 3852
4799 AAGAACCTTCTGAGCATCTCTTACGATCCCAAGGT-GCTTCCATGGCTCTTAAGAC 4857
3853 CCTTTGATTGGCCCAAAACAGGAAGTGGTTTCAAAGTGACAGAAACCAACACTTGTGTT 3912
4858 CCTTTGATTGGCCCAAAACAGGAAGTGGTTTCAAAGTGACAGAAACCAACACTTGTGTT 4917
3913 CAACCAAAATTTCAAGCAGTCTCCATCACTCAATCCAAATTCGATACCCAGCACTTTTGG 3972
4918 CAACCAAAATTTCAAGCAGTCTCCATCACTCAATCCAAATTCGATACCCAGCACTTTTGG 4977
3973 TTGCTCCAGATGTAGCACCTTTTATACCAAAACCGTGACGACGAGATTTGGTAGACTCCAG 4032
4978 TTGCTCCAGATGTAGCACCTTTTATACCAAAACCGTGACGACGAGATTTGGTAGACTCCAG 5037
4033 TTTGTTCTTATAGCTTCGGGAATAGACTTTTGGACGAGTACACAGGCCCCAACGAGT 4092
5038 TTTGTTCTTATAGCTTCGGGAATAGACTTTTGGACGAGTACACAGGCCCCAACGAGT 5097
4093 AATTAGAGAGTCAAGCACCACAAAGTAGTGAATAGACATCGGGCGGTTCAGTAGTCAAAAG 4152
5098 AATTAGAGAGTCAAGCACCACAAAGTAGTGAATAGACATCGGGCGGTTCAGTAGTCAAAAG 5157
4153 ACGCCAAACAAATTTCTGACAGGGAACCTTTTGTGACATCTTCAGAAAGTTCGTATTTCAG 4212
5158 ACGCCAAACAAATTTCTGACAGGGAACCTTTTGTGACATCTTCAGAAAGTTCGTATTTCAG 5217
4213 TAGTCAATTCGCCAGCATCAATATGGGATTTATACAGAAACCAACAGTGAAGTCAACAT 4272
5218 TAGTCAATTCGCCAGCATCAATATGGGATTTATACAGAAACCAACAGTGAAGTCAACAT 5277
4273 CTACCAACTTTGCGGTCTCAGAAAAAGCATAAACAGTTCCTACTACCGCCATTTAGTGAAC 4332
5278 CTACCAACTTTGCGGTCTCAGAAAAAGCATAAACAGTTCCTACTACCGCCATTTAGTGAAC 5337
4333 TTTTCAAAATCGCCAGTGGAGAGAAAGGCAAGGATCTAGCATTTAGCGGCAAGG 4392
5338 TTTTCAAAATCGCCAGTGGAGAGAAAGGCAAGGATCTAGCATTTAGCGGCAAGG 5397
4393 ATGCAACTTTTATCAACCCAGGCTCTATAGATAACCTTAGCGCTGGGATCATCTTTTGA 4452
5398 ATGCAACTTTTATCAACCCAGGCTCTATAGATAACCTTAGCGCTGGGATCATCTTTTGA 5457
4453 CAACCTTTCTGCAAAATCTAGGTCGCAAAATCACTTCATTGTATACCATTTGTGACACT 4512
5458 CAACCTTTCTGCAAAATCTAGGTCGCAAAATCACTTCATTGTATACCATTTGTGACACT 5517
4513 TGAGCAAGTTGTGATCAGCTCTCAAAATTTGGTTCCTCTGTAAACGAGTACTCAACTTGCA 4572
5518 TGAGCAAGTTGTGATCAGCTCTCAAAATTTGGTTCCTCTGTAAACGAGTACTCAACTTGCA 5577
4573 CATTAACCTGAAGCTCAGTTCAGTGAATTCAGTGAATTCAGGTTGTGAGTGTGAGTGTGAG 4632
5578 CATTAACCTGAAGCTCAGTTCAGTGAATTCAGTGAATTCAGTGTGAGTGTGAGTGTGAG 5637
4633 CATAGGGAACACAGGCTTTTCTTACCAACTCAAGGANTTATCAAACTCTGCAACACTTG 4692
5638 CATAGGGAACACAGGCTTTTCTTACCAACTCAAGGANTTATCAAACTCTGCAACACTTG 5697

QY	4693	CGTATGCAGGTACGACAGGAAATGTCTATCTTGAAGTCGACAGTCAGTGTAGTCTTGAG	4752
DB	5698	CGTATGCAGGTACGACAGGAAATGTCTATCTTGAAGTCGACAGTCAGTGTAGTCTTGAG	5757
QY	4753	AAATTCGAAGCCGTATTTTTTATTCAGTGTAGTCAGTCATCAGGAGATCCTCTCAGCCG	4812
DB	5758	AAATTCGAAGCCGTATTTTTTATTCAGTGTAGTCAGTCATCAGGAGATCCTCTCAGCCG	5817
QY	4813	GAGCATCTGGCGACCTCGCAGGTGGCATCACCGCGCCACAGTCGCGTTGCTGGCG	4872
DB	5818	GAGCATCTGGCGACCTCGCAGGTGGCATCACCGCGCCACAGTCGCGTTGCTGGCG	5877
QY	4873	CCTATATCGCGGACATCACCGATGGGAAAGATCGGCTCGCCATCTCGGCTCATGAGCG	4932
DB	5878	CCTATATCGCGGACATCACCGATGGGAAAGATCGGCTCGCCATCTCGGCTCATGAGCG	5937
QY	4933	CTTGTTTCGGGTGGGTATGGTGGCAGGCCCTGGTGGCGGGGACTGTTGGCGGCCATCT	4992
DB	5938	CTTGTTTCGGGTGGGTATGGTGGCAGGCCCTGGTGGCGGGGACTGTTGGCGGCCATCT	5997
QY	4993	CTTGTCATGCAC	5004
DB	5998	CTTTGGACCTGCAGGGGGGGGAAAGCCACGTGTGTCTCAAAATCTCTGATGTTA	6057
QY	5005	-----	5004
DB	6058	CATTGCACAAGATAAAATATATCATCATGAACAATAAAACTGTCTGTACATAAACAG	6117
QY	5005	-----	5004
DB	6118	TAATPACAAAGGGTGTATGAGCCATATTCACGGGAAACGCTTGTGTCGAAGCCGGATT	6177
QY	5005	-----	5004
DB	6178	AAATTCACAATGGATGCTGATTATATGGTATPAAATGGCTCGGATAAATGTCGGCA	6237
QY	5005	-----	5004
DB	6238	ATCAGTGCACAATCTATCGATTGTATGGAAAGCCGATGCCCCAGAGTTGTTCTGAA	6297
QY	5005	-----	5004
DB	6298	ACATGGCAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAACTGGCT	6357
QY	5005	-----	5004
DB	6358	GACGGAATTTATGCTCTCTCGACCATCAAGCAATTTTATCCGTACTCCTGATGATGCATG	6417
QY	5005	-----	5004
DB	6418	GTTACTCACCACCTGCGATCCCGCGGAAAAACAGCATTCAGGTATTAGAAATATCCTGA	6477
QY	5005	-----	5004
DB	6478	TTCAGGTCAAAATATTGTTGATGCTGCTGCGAGTGTCTCTGCGCGGTGCAATTCGATTCC	6537
QY	5005	-----	5004
DB	6538	TGTTTGTAAATTGCTTTTAAACAGCGATCGCGTATTTTCTGCTCGAGCGCAATCAGC	6597
QY	5005	-----	5004
DB	6598	AATGAATAACGGTTGGTTGATGCGAGTGATTTTGTACAGCGGTAAATGGCTGGCCTGT	6657
QY	5005	-----	5004
DB	6658	TGAACAAGTCTGGAAAGAAATGCATAAGCTTTTGGCATTTCTCACCGGATTCAGTCGTAC	6717
QY	5005	-----	5004
DB	6718	TCATGGTCAATTTCTCACTTGATTAACCTTATTTTGTACAGGGGAAATTAATAGTTGTAT	6777

QY	5005	-----	5000	-----
Db	6778	TGATGTTGGACGAGTCGGAAATCCGACACCGAGTACTTGCCATCCTATGGAACTG	6837	
QY	5005	-----	5004	-----
Db	6838	CCTCGGTGAGTTTTCTCCTTCATTACAGAAAAGCGCTTTTTCAAAAATATGGTATTGATAA	6897	
QY	5005	-----	5004	-----
Db	6898	TCCTGATATGAATAAAATTGACGTTTCATTATTGATGCTCGATGAGTTTTTTCTAAATCAGAAATT	6957	
QY	5005	-----	5004	-----
Db	6958	GGTTAAATGGTTGTAAACATCTGGCAGAGCAATTACGTCGACTTTGACGGGACGGCGGCTTTGT	7017	
QY	5005	-----	5004	-----
Db	7018	TGAATAAATCGAACTTTTGTGCTGAGTTGAAGGATCAGATCAGCATCTTCCGACCAAGCA	7077	
QY	5005	-----	5004	-----
Db	7078	GACCGTTCCGTGGCAAGCAAAAGTTCAAATCACCAACTGGTCCACCTACAAACAAGCT	7137	
QY	5005	-----	5004	-----
Db	7138	CTCATCAACCGTGGCTCCCTCACTTTCTGCTGGATGATGGGGCGATTCAGGCGCTGGTAT	7197	
QY	5005	-----	5012	-----CATTCCTT
Db	7198	GAGTCAGACAAACCTTCTTCAGAGGACAGACCTCAGCGCCCCCCCCCTGCAGGTCCTC	7257	
QY	5013	CGCGCGGGTGCTCAACGGCCTCAACCTACTACTGGGCTGCTTCTCTAAATGCAGGAGTCG	5072	
Db	7258	ACGCGCGGGTGCTCAACGGCCTCAACCTACTACTGGGCTGCTTCTCTAAATGCAGGAGTCG	7317	
QY	5073	CATAAGGGAGAGCGTCGAGTATCTATGATTTGGAAGTATGGAAATGGGATACCCGCAATTC	5132	
Db	7318	CATAAGGGAGAGCGTCGAGTATCTATGATTTGGAAGTATGGAAATGGGATACCCGCAATTC	7377	
QY	5133	TTCAAGTCTTTGAGGTCTCTATCAGAGTATGCCCACTAAAGCAACCGGAGGAGGAGAT	5192	
Db	7378	TTCAAGTCTTTGAGGTCTCTATCAGAGTATGCCCACTAAAGCAACCGGAGGAGGAGAT	7437	
QY	5193	TTCAATGTTAAATTTCTCTGACTTTTGTGCTATCAGTAGACTCGAACTGTGAGACTATCTCG	5252	
Db	7438	TTCAATGTTAAATTTCTCTGACTTTTGTGCTATCAGTAGACTCGAACTGTGAGACTATCTCG	7497	
QY	5253	GTTATGACAGCAGAAATGTCTCTTCCTGGAGACAGTAAATGAACTCCCAACCAATAAGAAA	5312	
Db	7498	GTTATGACAGCAGAAATGTCTCTTCCTGGAGACAGTAAATGAACTCCCAACCAATAAGAAA	7557	
QY	5313	TCCCTGTTATCAGGAAACAACTTCTGTTTTCGAACTTTTCGGTGCCTTGAACCTATAAAA	5372	
Db	7558	TCCCTGTTATCAGGAAACAACTTCTGTTTTCGAACTTTTCGGTGCCTTGAACCTATAAAA	7617	
QY	5373	TGTAGAGTGGATATGTCCGGTAGGAATGGAGCGGCAAAATGCTTACCTTCTGACCTTCA	5432	
Db	7618	TGTAGAGTGGATATGTCCGGTAGGAATGGAGCGGCAAAATGCTTACCTTCTGACCTTCA	7677	
QY	5433	AGAGGTATGTAGGGTTTGTAGATACGTATGCCAACTTCAGTGACAAAGTTCGTATTTCGT	5492	
Db	7678	AGAGGTATGTAGGGTTTGTAGATACGTATGCCAACTTCAGTGACAAAGTTCGTATTTCGT	7737	
QY	5493	TCAAACCAATCCGAATCCAGAGAAATCAAAGTTGTTGTCTACTATTGATCCCAAGCCAGT	5552	
Db	7738	TCAAACCAATCCGAATCCAGAGAAATCAAAGTTGTTGTCTACTATTGATCCCAAGCCAGT	7797	
QY	5553	CGGCTTTGAAACTGCAAAATAGTGTCTGTTTTCAGGTTCATCTTTGTATGATTAAT	5612	
Db	7798	CGGCTTTGAAACTGCAAAATAGTGTCTGTTTTCAGGTTCATCTTTGTATGATTAAT	7857	
QY	5613	CTAGCTTTTGTATCTAAATTAATCTTTGACGAGCCAAAGGGATAAATACCCAAATCTAAAACT	5672	

7858 CTAGCTTTGATCTAAATAATCTTGACGAGCCAGCGGATAAATACCCAAATCTAAACT 7917
5673 CTTTTAAACGTTAAAGAGCAAGTATGTCTGCTGTTATTAACCCCAAAATCAGCTCGTA 5732
7918 CTTTTAAACGTTAAAGAGCAAGTATGTCTGCTGTTATTAACCCCAAAATCAGCTCGTA 7977
5733 GTCTGATCTCATCAACTTGAGGGGCACTATCTGTTTAGAGAAATTTGCGGAGATCG 5792
7978 GTCTGATCTCATCAACTTGAGGGGCACTATCTGTTTAGAGAAATTTGCGGAGATCG 8037
5793 ATATCGAGAAAAGGTAGCTGATTTTAAACGTGAAATTTATCTCAAGATCTCTGCTCG 5852
8038 ATATCGAGAAAAGGTAGCTGATTTTAAACGTGAAATTTATCTCAAGATCTCTGCTCG 8097
5853 CCGCTTTGCTGATGACGCTGAAACCTCTGACACATGACAGTCCCGGAGAGCGTACAG 5912
8098 CCGCTTTGCTGATGACGCTGAAACCTCTGACACATGACAGTCCCGGAGAGCGTACAG 8157
5913 CTTGCTGTAAGCGGATCGGGAGCAGACAAGCCCGTCAGGGCCGTCAGCGGTGTTG 5972
8158 CTTGCTGTAAGCGGATCGGGAGCAGACAAGCCCGTCAGGGCCGTCAGCGGTGTTG 8217
5973 GCGGCTGTCGGGCGCAGCCATGACCCAGTCAGCTACGTAGCGATAGCGAGTGTATCTGCT 6032
8218 GCGGCTGTCGGGCGCAGCCATGACCCAGTCAGCTACGTAGCGATAGCGAGTGTATCTGCT 8277
6033 TAACTATGCGGATCAGAGCAGATTTGTAAGAGTGCACCATATGCGGTGTGAAATACC 6092
8278 TAACTATGCGGATCAGAGCAGATTTGTAAGAGTGCACCATATGCGGTGTGAAATACC 8337
6093 GCACAGATCGTAAGAGAAATACCGGATCAGGGCTCTTCGCTTCTCTGCTCACTGA 6152
8338 GCACAGATCGTAAGAGAAATACCGGATCAGGGCTCTTCGCTTCTCTGCTCACTGA 8397
6153 CTCGCTGCGCTCGCTGCTGCTGCGGAGCGGTATCAGCTCACTCAAAAGGCGTAAT 6212
8398 CTCGCTGCGCTCGCTGCTGCGGAGCGGTATCAGCTCACTCAAAAGGCGTAAT 8457
6213 ACGGTTATCCAGAAATCAGGGATACGAGAGAAAGAACATGTGAGCAAAAGGCCAGCA 6272
8458 ACGGTTATCCAGAAATCAGGGATACGAGAGAAAGAACATGTGAGCAAAAGGCCAGCA 8517
6273 AAAGGCCAGAACCTTAAAGGCGCGTTCGCTGCGGTTTTCATAGGCTCCGCCCCC 6332
8518 AAAGGCCAGAACCTTAAAGGCGCGTTCGCTGCGGTTTTCATAGGCTCCGCCCCC 8577
6333 TGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGGAAACCCGACAGACTATA 6392
8578 TGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGGAAACCCGACAGACTATA 8637
6393 AAGATACAGGGGTTTCCCGCTGGAAGTCCCTCGTGGCTCTCTGTTCCGACCTGCC 6452
8638 AAGATACAGGGGTTTCCCGCTGGAAGTCCCTCGTGGCTCTCTGTTCCGACCTGCC 8697
6453 GTTACCGGATACCTGTCGCGCTTCTCCCTTCGGAAGCGTGGCGCTTCTCATAGCTC 6512
8698 GTTACCGGATACCTGTCGCGCTTCTCCCTTCGGAAGCGTGGCGCTTCTCATAGCTC 8757
6513 ACGTGTAGGATCTCAGTTCCGCTGAGTGGTCCCTCGTGGCTCTCTGTTCCGACCTGCC 6572
8758 ACGTGTAGGATCTCAGTTCCGCTGAGTGGTCCCTCGTGGCTCTCTGTTCCGACCTGCC 8817
6573 ACCCCCGTTACGCCCGCTGCGCTTATCCGCTAACCTATCGCTTGGTCCGACCTGCC 6632
8818 ACCCCCGTTACGCCCGCTGCGCTTATCCGCTAACCTATCGCTTGGTCCGACCTGCC 8877
6633 GGTAAAGACAGCTTATCGCCACTGCGAGCAGCCACTGTTAACAGGATTAGCAGCGAG 6692
8878 GGTAAAGACAGCTTATCGCCACTGCGAGCAGCCACTGTTAACAGGATTAGCAGCGAG 8937
6693 GTATGAGCGGTGCTACAGATTTCTGAAAGTGGTGGCTTAACCTAGCGCTACATAGAG 6752

8938 GTATGAGCGGTCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAG 8997
6753 GACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGGTAG 6812
8998 GACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGGTAG 9057
6813 CTCTTGATCCGGCAAAACAAACACCCGCTGCTAGCGGTGGTGTGTTTGTGTTGCAAGCAGCA 6872
9058 CTCTTGATCCGGCAAAACAAACACCCGCTGCTAGCGGTGGTGTGTTTGTGTTGCAAGCAGCA 9117
6873 GATTACGCGCAGAAAAAAGGATCTCAAGAGATCTCTTGATCTTTTCTACGGGGTCTGA 6932
9118 GATTACGCGCAGAAAAAAGGATCTCAAGAGATCTCTTGATCTTTTCTACGGGGTCTGA 9177
6933 CGCTCAGTGAACGAAACCTCAGTTTAAAGGATTTTGGTCATGAGATTAATCAAAAAGGAT 6992
9178 CGCTCAGTGAACGAAACCTCAGTTTAAAGGATTTTGGTCATGAGATTAATCAAAAAGGAT 9237
6993 CTTACACCTAGATCTTTTAAATTAAGAGATTTTAAATCAATCTAAAGTATATATGA 7052
9238 CTTACACCTAGATCTTTTAAATTAAGAGATTTTAAATCAATCTAAAGTATATATGA 9297
7053 GTAAACTTGTCTGACAGTTTACCAATGCTTAATCAGTGAAGGACCTATCTCAGCGATCTG 7112
9298 GTAAACTTGTCTGACAGTTTACCAATGCTTAATCAGTGAAGGACCTATCTCAGCGATCTG 9357
7113 TCTATTTGCTTCATCCATAGTTCCCTGACTCCCGCTGCTGTAGATAAATCAAGTACGGGA 7172
9358 TCTATTTGCTTCATCCATAGTTCCCTGACTCCCGCTGCTGTAGATAAATCAAGTACGGGA 9417
7173 GGGTTACCATCTGGCCCCAGTCTGCAATGATACCGCAGAGCCAGCTCAGCGGCTCC 7232
9418 GGGTTACCATCTGGCCCCAGTCTGCAATGATACCGCAGAGACCCAGCTCAGCGGCTCC 9477
7233 AGATTTATCAGCAATAAACCAGCAGCCGGAAGGCGCGAGCGAGAGTGGTCTGCAAC 7292
9478 AGATTTATCAGCAATAAACCAGCAGCCGGAAGGCGCGAGCGAGAGTGGTCTGCAAC 9537
7293 TTTATCGGCTCCCATCAGTCTAATTTGCTGCGGGAGCTAGAGTAAAGTATGCGCC 7352
9538 TTTATCGGCTCCCATCAGTCTAATTTGCTGCGGGAGCTAGAGTAAAGTATGCGCC 9597
7353 AGTTAATAGTTTGGCAACGTTGTCATTGCTGAGGCGATCGTGGTGTACGCTCGTC 7412
9598 AGTTAATAGTTTGGCAACGTTGTCATTGCTGAGGCGATCGTGGTGTACGCTCGTC 9657
7413 GTTGTATGCTTCATTCAGCTCCGTTCCCAACGATCAAGCGAGTTTACATGATCCCC 7472
9658 GTTGTATGCTTCATTCAGCTCCGTTCCCAACGATCAAGCGAGTTTACATGATCCCC 9717
7473 CATGTTGTGCAAAAAGCGTTAGCTCTTCGCTCCGATCGTTGTGTCAGAAAGTAAATT 7532
9718 CATGTTGTGCAAAAAGCGTTAGCTCTTCGCTCCGATCGTTGTGTCAGAAAGTAAATT 9777
7533 GCGCGCAGTGTATCACTCATGTTTATGGCAGCACTGCATAATTTCTTACTGTCTATGCC 7592
9778 GCGCGCAGTGTATCACTCATGTTTATGGCAGCACTGCATAATTTCTTACTGTCTATGCC 9837
7593 ATCCGTAAAGATGCTTTTCTGTGATGCTGAGTACTCAACCAAGTCTATCTGAGAAATAGTG 7652
9838 ATCCGTAAAGATGCTTTTCTGTGATGCTGAGTACTCAACCAAGTCTATCTGAGAAATAGTG 9897
7653 TATGCGCGCAGCGAGTTGCTTTCGCGCGCTCAACAGCGGATAATACCGCGCCACATAG 7712
9898 TATGCGCGCAGCGAGTTGCTTTCGCGCGCTCAACAGCGGATAATACCGCGCCACATAG 9957
7713 CAGAACTTTAAAGTGTCTCATCTATGAAACCGTTCTTGGGCGGAAAACTCTCAAGGAT 7772
9958 CAGAACTTTAAAGTGTCTCATCTATGAAACCGTTCTTGGGCGGAAAACTCTCAAGGAT 10017
7773 CTTACCGCTCTGAGATCCAGTTTCGATGTAAACCCACTCGTGCACCAACTGATCTTCAGC 7832
10018 CTTACCGCTCTGAGATCCAGTTTCGATGTAAACCCACTCGTGCACCAACTGATCTTCAGC 10077

QY 1253 -----GTATCATGGCCCCCGGAGAGGATAAATTTGGCTTCCCGGTGTC 1299
Db 1381 ACATCATATTAAAGGATGAAGAACACCGGAGCGGACAGGCTGATGACTCAGAGGTGA 1440
QY 1300 A----- 1300
Db 1441 AGGACAAGTTGAACGCTTTGGCCATCTCGGTGATGAACCACTGGCCAGGAGTGAACCTGC 1500
QY 1301 ----- 1300
Db 1501 GGGTGACCGAGGGCTGGGACGNAGATGCCACCCTCAGAGGAGTCTCTGCCTACGAGG 1560
QY 1301 ----- 1300
Db 1561 GCCCGCAGTGGACATCACCACTGTGACCGGACCGCAGCAAGTACGGCATGTCTGGCCC 1620
QY 1301 ----- 1300
Db 1621 GCCTGGCGGTGGAGCCCGGCTTCGACTGGGTGTACTACGATGCCAAGSCACATATCCACT 1680
QY 1301 ----- 1300
Db 1681 GCTCGGTGAAGACAGAGAAGTCCGGTGGCGGCCAAATCGGGAGGGCTGCACCCAGAGGGC 1740
QY 1301 ----- 1300
Db 1741 CCACAATCAAGCCCTGTCTCCATGCAAAATGCCAGCACCTAACCTCTTGGGTGGACCAT 1800
QY 1301 ----- 1300
Db 1801 CCGTCTTTCATCTTCCCTCCAAAGATCAAGGATGTACTCATGATCTCCCTGAGCCCCATAG 1860
QY 1301 ----- 1300
Db 1861 TCACATGTGTGGTGGATGTGAGCGAGGATGACCCAGATGTCACAGATCCTCCAGTGGTTG 1920
QY 1301 ----- 1300
Db 1921 TGAACAACGTGGAAGTACACACAGCTCAGACACAAACCCATAGAGAGGATTTACCAAGTA 1980
QY 1301 ----- 1300
Db 1981 CACTTCGGTGTGTGAGTCCCTCCATCCAGCACAGGACTGGATGAGTGGCAAGGAGT 2040
QY 1301 ----- 1300
Db 2041 TCAATGCAAGTCAACAACAAGACCTCCAGCGCCCATCGAGAGAACCATCTCAAAAC 2100
QY 1301 ----- 1300
Db 2101 CCAAGGGTCAGTAAGAGCTCCACAGGTATATGTCTTCCCTCCACCAGAGAGATGA 2160
QY 1301 -----CGGCCAGAGTGCACGGAGAGAGGTTGCTGTTTGTATGACAGTGTCC 1348
Db 2161 CTAAGAAACAGGTCACTCTGACCTGCTGATGACAGACTTCATGCTGAGACATTTACG 2220
QY 1349 GGG----- 1351
Db 2221 TGGAGTGGACCAACAAACGGGAAAAACAGAGCTAACTACAGAACACTGAAACAGTCTCTGG 2280
QY 1352 -----GATTCGGTGGTGTCTCCACCCCATGGCCATCGACCAACACTCAAG 1396
Db 2281 ACTCTGATGGTCTTACTTCTATGTCACAGCAAGCTGAGAGTGGAAAGAAAGAACTGGGTGG 2340
QY 1397 AAGAGAAATGTCCTCTTAACTAGTGG----- 1423
Db 2341 AAAGAAATAGTACTCTCTGTTTCAGTGGTCCACGAGGGTCTGCACAAATCACACAGACTA 2400
QY 1424 -----CGTGAATTC 1433
Db 2401 AGAGCTTCTCCGAGTCCGGGTAAATGAGCTCAGATCGATTCATGGATCTCTCATCTCC 2460

QY 1434 CCTAGGCGCGCCGCAATTAATTCGCTTAGACATGACTGTCTTCCCTCAGTTCAGTTGGGC 1493
Db 2461 CAATCGCGCGCCGCAATTAATTCGCTTAGACATGACTGTCTTCCCTCAGTTCAGTTGGGC 2520
QY 1494 ACTTCAGAGAGACCGGCTCTTCTGATTTCTAATCAAGAGGATGTGAGATGCCATTTCG 1553
Db 2521 ACTTCAGAGAGACCGGCTCTTCTGATTTCTAATCAAGAGGATGTGAGATGCCATTTCG 2580
QY 1554 CTGAGAGATGACGCTTCATTTTGTACTTTTTTATTTTAACTATATAGTATAGGAT 1613
Db 2581 CTGAGAGATGACGCTTCATTTTGTACTTTTTTATTTTAACTATATAGTATAGGAT 2640
QY 1614 TTTTTCGTCAATTTTCTCTCTGACGAGTCTCTCTGATGAGCTTCTCTGATGAGCTTCTCCAGC 1673
Db 2641 TTTTTCGTCAATTTTCTCTCTGACGAGTCTCTCTGATGAGCTTCTCTGATGAGCTTCTCCAGC 2700
QY 1674 TGATGAATATCTTGTGTAGGGGTTTGGCAAAATCAATCGAGTTTCATGTTTTCTTTCTTGT 1733
Db 2701 TGATGAATATCTTGTGTAGGGGTTTGGCAAAATCAATCGAGTTTCATGTTTTCTTTCTTGT 2760
QY 1734 ATTTCACATCTCTCTCTCAGAGTACAGAGATTAAGTGAGAGTTCGTTTGTGCAAGCTTA 1793
Db 2761 ATTTCACATCTCTCTCTCAGAGTACAGAGATTAAGTGAGAGTTCGTTTGTGCAAGCTTA 2820
QY 1794 TCGATAGCTTTAATCGGTAGTTTATCACAGTTAAATTTGCTAAACAGTTCAGGACCCGT 1853
Db 2821 TCGATAGCTTTAATCGGTAGTTTATCACAGTTAAATTTGCTAAACAGTTCAGGACCCGT 2880
QY 1854 GTATGAATCTAAATGCGCTCATCTCTCGGCAACCGTCAACCTGGATGCTGTAG 1913
Db 2881 GTATGAATCTAAATGCGCTCATCTCTCGGCAACCGTCAACCTGGATGCTGTAG 2940
QY 1914 GCATAGCTTGGTTATGCCGTACTCGCGGCTCTTGGGGATATCGTTCATTCGACA 1973
Db 2941 GCATAGCTTGGTTATGCCGTACTCGCGGCTCTTGGGGATATCGTTCATTCGACA 3000
QY 1974 GCATCGCAGTCACTATGCGTCTCTAGCGTATATGCGTTGTATGCAATTTCTATGCG 2033
Db 3001 GCATCGCAGTCACTATGCGTCTCTAGCGTATATGCGTTGTATGCAATTTCTATGCG 3060
QY 2034 CACCGTTCTCGGAGACTGTCCGACCGTTTGGCGCGCCCGCAGTCTCTGCTGCTTCCG 2093
Db 3061 CACCGTTCTCGGAGACTGTCCGACCGTTTGGCGCGCCCGCAGTCTCTGCTGCTTCCG 3120
QY 2094 TACTTCGAGCACTATCGACTAGCGATCATGCGGACCAACCCGCTCTGTGTGATCTATC 2153
Db 3121 TACTTCGAGCACTATCGACTAGCGATCATGCGGACCAACCCGCTCTGTGTGATCTATC 3180
QY 2154 GAATCTAAATGTAAGTTAAATCTCTAAATAATTAATAAGTCCAGTTTCTCCATACGA 2213
Db 3181 GAATCTAAATGTAAGTTAAATCTCTAAATAATTAATAAGTCCAGTTTCTCCATACGA 3240
QY 2214 ACCTTAACAGATTCGGGTGAGCTCTAGACCTTCAACGAGCCAGAGATCCATCTGCT 2273
Db 3241 ACCTTAACAGATTCGGGTGAGCTCTAGACCTTCAACGAGCCAGAGATCCATCTGCT 3300
QY 2274 TGGCCAAATATGTTTCAGTCCCTCAGGATTAACGTTCTGTGAGTGAATCTCTGGAAG 2333
Db 3301 TGGCCAAATATGTTTCAGTCCCTCAGGATTAACGTTCTGTGAGTGAATCTCTGGAAG 3360
QY 2334 GTTTCAGTGTAACTCGCTGTATTGACGGGCAATATCCGTAGCTTGGGAAAAGTGTGGTTG 2393
Db 3361 GTTTCAGTGTAACTCGCTGTATTGACGGGCAATATCCGTAGCTTGGGAAAAGTGTGGTTG 3420
QY 2394 GTACCGGAGGATAATCTCCACAACCTCTCTGGAGAGTAGGCACCAACAACACAGATCCA 2453
Db 3421 GTACCGGAGGATAATCTCCACAACCTCTCTGGAGAGTAGGCACCAACAACACAGATCCA 3480
QY 2454 GCCTGTGTACTTGTATCAACATAAAGAGAGCAATCTCGAATTTGCAAGGATCAAGTGTCA 2513
Db 3481 GCCTGTGTACTTGTATCAACATAAAGAGAGCAATCTCGAATTTGCAAGGATCAAGTGTCA 3540
QY 2514 GGAGCGTACTGATTGGACATTTCCAAAGCCTGCTCTAGTTTCCAAACCGATAGGGTTGTA 2573

QY 4733 ACAGTGAGTGTAGTCTTCTGAGAAAATCTCGAAGCCGTTATTTTATATATATAGTCAAGTCA 4792
| | | | |
Db 5759 ACAGTGAGTGTAGTCTTCTGAGAAAATCTCGAAGCCGTTATTTTATATATAGTCAAGTCA 5818
| | | | |
QY 4793 TCAGGAGATCCTCTACGCCGACGATCGTGGCCGACCTGACAGTCCGATCACCAGCCG 4852
| | | | |
Db 5819 TCAGGAGATCCTCTACGCCGACGATCGTGGCCGACCTGACAGTCCGATCACCAGCCG 5878
| | | | |
QY 4853 CACAGTGCGGTGTGTGGCGCTATATCCCGACATCAACCATCGGGAAGATCGGGCTCG 4912
| | | | |
Db 5879 CACAGTGCGGTGTGTGGCGCTATATCCCGACATCAACCATCGGGAAGATCGGGCTCG 5938
| | | | |
QY 4913 CCATCTCGGCTCATGCGCTGTGTTTGGCGGTGTGTTGTCAGCCGCTGCGCGG 4972
| | | | |
Db 5939 CCATCTCGGCTCATGCGCTGTGTTTGGCGGTGTGTTGTCAGCCGCTGCGCGG 5998
| | | | |
QY 4973 GGGACTGTGGCGCATCTCCTTCATGAC- 5004
| | | | |
Db 5999 GGGACTGTGGCGCATCTCCTTCATGAC- 6058
| | | | |
QY 5005 - 5004
| | | | |
Db 6059 TCTCAAAATCTGTATGTACATTCGACAGATAAAAATATATCATCATGAACAATAAAA 6118
| | | | |
QY 5005 - 5004
| | | | |
Db 6119 CTGTCTGCTTACATAAAACAGTAATACAGGGGTGTATGAGCCATATTCACAGGGAACG 6178
| | | | |
QY 5005 - 5004
| | | | |
Db 6179 TCTTCTCAAGCCCGGATTAATTTCCAAATGATGATGCTGATTTATATGGGTATAAATGG 6238
| | | | |
QY 5005 - 5004
| | | | |
Db 6239 GCTCCGATAATGTCCGGCAATCAGTCCGACAACTATATCATGATGTATGGGAAGCCCGAT 6298
| | | | |
QY 5005 - 5004
| | | | |
Db 6299 GCGCAGAGTGTCTTCTGAACATCGCAAGTAGCGTTGCCATGATGTTACAGATGAG 6358
| | | | |
QY 5005 - 5004
| | | | |
Db 6359 ATGTGACATAAATGCTGACGGAATTTATGCTCTTCCGACCATCAAGCATTTTATC 6418
| | | | |
QY 5005 - 5004
| | | | |
Db 6419 CGTACTCTGATGATGATGTTACTCACCAGTCCGATCCCGGGAACAGCATTCAG 6478
| | | | |
QY 5005 - 5004
| | | | |
Db 6479 GTATTAGAAGATATCCTGATTAGGTGAAAATATTTGTTGATGCGCTGGCAGTGTCTCG 6538
| | | | |
QY 5005 - 5004
| | | | |
Db 6539 GCGCGTTGCATTCGATTCCTGTTTGTGTAATGTTGCTTTTAAACAGGATCGCGTATTTTCT 6598
| | | | |
QY 5005 - 5004
| | | | |
Db 6599 CTCGCTCAGCGCAATCAGGAATGAATAACGGTTTGGTTGATGCGAGTATTTTGTATGAC 6658
| | | | |
QY 5005 - 5004
| | | | |
Db 6659 GAGCGTAATGGTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAGCTTTTGGCCATTC 6718
| | | | |
QY 5005 - 5004
| | | | |
Db 6719 TCACCGGATTCAGTCGCTCACTCATGGTATTTCTCACTTGTATTAACCTTATTTTTCAGCAG 6778
| | | | |
QY 5005 - 5004
| | | | |
Db 6779 GGGAAATTAATAGTGTATTGATGTGGAAGAGTCCGGAATCGCAGACCGGATACCGAGAT 6838
| | | | |

QY 5005 - 5004
| | | | |
Db 6839 CTTGCCATCCTATGGAACCTCCTCGGTGAGTTTTCTCTTCAATTACAGAAACGGTTTTT 6898
| | | | |
QY 5005 - 5004
| | | | |
Db 6899 CAAAAATATGGTATTGATTAATCCTGATATGAATAAATTCAGATTTTCATTTGATGCTCGAT 6958
| | | | |
QY 5005 - 5004
| | | | |
Db 6959 GAGTTTTTCTTAATCAGAAATTTGGTTAATTTGGTTGTAAACACTGGCAGAGCATTTACGCTGACT 7018
| | | | |
QY 5005 - 5004
| | | | |
Db 7019 TGACGGACGCGCGCTTTGTTGAATAAATCGAACTTTTCTGCTGAGTTGAAGGATCAGATCA 7078
| | | | |
QY 5005 - 5004
| | | | |
Db 7079 CGCATCTTCCGACACAGCAGACCGTTCCGTGSCAAAGCAAAAGTTCAAAATCACCACACT 7138
| | | | |
QY 5005 - 5004
| | | | |
Db 7139 GTTCCACCTACAAACAAAGCTCTCATCAACCGTGGTCCCTCACTTTCTGGCTGGATGATG 7198
| | | | |
QY 5005 - 5004
| | | | |
Db 7199 GGGCGATTACGGCCTGTATGATGATCAGCAACACCTTTCTTACAGAGCAGACCTCAGCGCC 7258
| | | | |
QY 5005 - 5004
| | | | |
Db 7259 CCCCCCCTGTCAGGTCCACGCGCGGTGCTCAACGCGCTCAACCTACTACTCGGGCT 7318
| | | | |
QY 5053 GCTTCTTAATGTCAGGAGTGCATAAGGAGAGCGTGCAGTATCTATGATTGGAAGTATGG 5112
| | | | |
Db 7319 GCTTCTTAATGTCAGGAGTGCATAAGGAGAGCGTGCAGTATCTATGATTGGAAGTATGG 7378
| | | | |
QY 5113 GAATGTGATPACCGCATCTTCAGTGTCTTGAGGTCTCTATCATAGATTATGCCCAACTA 5172
| | | | |
Db 7379 GAATGTGATPACCGCATCTTCAGTGTCTTGAGGTCTCTATCATAGATTATGCCCAACTA 7438
| | | | |
QY 5173 AAGCAACCGGAGGAGGAGATTTTCATGTAATTTCTCTGACTTTTGGTGCATCAGTAGACT 5232
| | | | |
Db 7439 AAGCAACCGGAGGAGGAGATTTTCATGTAATTTCTCTGACTTTTGGTGCATCAGTAGACT 7498
| | | | |
QY 5233 CGAACTGTGAGACTATCTCGGTTATGACAGCAGAAATGTCTTCTTGGAGACAGTAAATG 5292
| | | | |
Db 7499 CGAACTGTGAGACTATCTCGGTTATGACAGCAGAAATGTCTTCTTGGAGACAGTAAATG 7558
| | | | |
QY 5293 AAGTCCCAACCAATAAAGAAATCTTGTATCAGGAACAAACTTCTTGTTCGAACCTTTT 5352
| | | | |
Db 7559 AAGTCCCAACCAATAAAGAAATCTTGTATCAGGAACAAACTTCTTGTTCGAACCTTTT 7618
| | | | |
QY 5353 CGGTCTTGAACATAAATAATGATAGTGTGATATGTCGGGTAGGAATGAGCGGCAAT 5412
| | | | |
Db 7619 CGGTCTTGAACATAAATAATGATAGTGTGATATGTCGGGTAGGAATGAGCGGCAAT 7678
| | | | |
QY 5413 GCTTACCTTCTGACCTTCAAGAGGTATGTAGGTTTGTAGATPACTGATGCCAATTCAG 5472
| | | | |
Db 7679 GCTTACCTTCTGACCTTCAAGAGGTATGTAGGTTTGTAGATPACTGATGCCAATTCAG 7738
| | | | |
QY 5473 TGACAAAGTGTCTTATTTCTGTAACCAATCCGAATCCAGAGAAATCAAAAGTTTGTTC 5532
| | | | |
Db 7739 TGACAAAGTGTCTTATTTCTGTAACCAATCCGAATCCAGAGAAATCAAAAGTTTGTTC 7798
| | | | |
QY 5533 TACTATTGATCCAAAGCCAGTGGCGTCTTGAAACTGACAAATAGTGTCTCGTGTTCGAGG 5592
| | | | |
Db 7799 TACTATTGATCCAAAGCCAGTGGCGTCTTGAAACTGACAAATAGTGTCTCGTGTTCGAGG 7858
| | | | |
QY 5593 TCATCTTTGTATGAATAAATCTAGCTTTTGTATCTAAATAATCTTGAAGCGCAAGCGAT 5652
| | | | |
Db 7859 TCATCTTTGTATGAATAAATCTAGCTTTTGTATCTAAATAATCTTGAAGCGCAAGCGAT 7918
| | | | |
QY 5653 AAATACCCAAATCTAAATACTCTTTTAAACGTTAAAGGACAAATGATGTCTGCTCTATT 5712
| | | | |

Db 7919 AATACCCAAATCTAAACCTCTTTTAAAGCTTAAAGGCAAGTATGCTGCTGTATT 7978
Qy 5713 AAACCCCAATCAGCTCGTAGTCGATCCCTCATCAACTTGAGGGGCACTATCTTGTTTAA 5772
Db 7979 AAACCCCAATCAGCTCGTAGTCGATCCCTCATCAACTTGAGGGGCACTATCTTGTTTAA 8038
Qy 5773 GAGAAATTTGCGGAGATCGATATCGAGAAAGGTTACGCTGATTTTAAACGTGAATTT 5832
Db 8039 GAGAAATTTGCGGAGATCGATATCGAGAAAGGTTACGCTGATTTTAAACGTGAATTT 8098
Qy 5833 ATCTCAAGATCTCTGCTCGCGCTTTTCGGTGTAGACGGTGAATACTTGACACATGCA 5892
Db 8099 ATCTCAAGATCTCTGCTCGCGCTTTTCGGTGTAGACGGTGAATACTTGACACATGCA 8158
Qy 5893 GCTCCCGAGACGGTCAAGCTTGTCTGTAAAGCGGATCCGGGAGCAGACGCGCTCA 5952
Db 8159 GCTCCCGAGACGGTCAAGCTTGTCTGTAAAGCGGATCCGGGAGCAGACGCGCTCA 8218
Qy 5953 GCGCGCTCAGCGGCTGTGCGGCTGTCCGGGCGCAGCCATGACCCAGTCAAGTACGCA 6012
Db 8219 GCGCGCTCAGCGGCTGTGCGGCTGTCCGGGCGCAGCCATGACCCAGTCAAGTACGCA 8278
Qy 6013 TAGCGGAGTATATCTGGCTTAATATGCGGCATCAGAGCAGATTTGACGAGTGCAC 6072
Db 8279 TAGCGGAGTATATCTGGCTTAATATGCGGCATCAGAGCAGATTTGACGAGTGCAC 8338
Qy 6073 CATATCGGCTGAATACCCACAGATCGTAAAGGAGAAATACCGCATCAGCGCTCT 6132
Db 8339 CATATCGGCTGAATACCCACAGATCGTAAAGGAGAAATACCGCATCAGCGCTCT 8398
Qy 6133 TCCGCTTCCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 6192
Db 8399 TCCGCTTCCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 8458
Qy 6193 GCTCACTAAAGGGGTAATACGGTTATCCACAGATCAGGGGATACGAGGAGAAAC 6252
Db 8459 GCTCACTAAAGGGGTAATACGGTTATCCACAGATCAGGGGATACGAGGAGAAAC 8518
Qy 6253 ATGTGAGCAAAAGGCCAGAAAGCCAGAAACCGTAAAGGCGCGGTTGCTGGCGTTT 6312
Db 8519 ATGTGAGCAAAAGGCCAGAAAGCCAGAAACCGTAAAGGCGCGGTTGCTGGCGTTT 8578
Qy 6313 TTCCATAGGCTCCGCCCCCTTGAGAGCATCAAAAATCGACCTCAAGTCAGAGTGG 6372
Db 8579 TTCCATAGGCTCCGCCCCCTTGAGAGCATCAAAAATCGACCTCAAGTCAGAGTGG 8638
Qy 6373 CGAAACCCGACAGGACTATAAGATACAGGCGTTTCCCTCGGAAGCTCCCTCGTCGC 6432
Db 8639 CGAAACCCGACAGGACTATAAGATACAGGCGTTTCCCTCGGAAGCTCCCTCGTCGC 8698
Qy 6433 TCTCCTGTTCGACCCCTCCGCTTACCGGATACCTGTCCGCTTTTCCCTTCGGGAAGC 6492
Db 8699 TCTCCTGTTCGACCCCTCCGCTTACCGGATACCTGTCCGCTTTTCCCTTCGGGAAGC 8758
Qy 6493 GTGGCGCTTCTCATAGCTCAGCTGTAGTATCTCAGTTCGGGTAGGTCGTTTCGCTCC 6552
Db 8759 GTGGCGCTTCTCAGTGTAGTATCTCAGTTCGGGTAGGTCGTTTCGCTCC 8818
Qy 6553 AAGCTGGGCTGTGTCAGCAAAACCCCGTTTACCGGATACCTGTCCGCTTTTCCCTTCGGGAAGC 6612
Db 8819 AAGCTGGGCTGTGTCAGCAAAACCCCGTTTACCGGATACCTGTCCGCTTTTCCCTTCGGGAAGC 8878
Qy 6613 TATCGTCTTGAGTCAACCCCGTAAAGACAGCACTTATCGCCACTTGGCAGGACCACTGCT 6672
Db 8879 TATCGTCTTGAGTCAACCCCGTAAAGACAGCACTTATCGCCACTTGGCAGGACCACTGCT 8938
Qy 6673 AACAGGATTTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTCTTTGAAGTGTGGCT 6732
Db 8939 AACAGGATTTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTCTTTGAAGTGTGGCT 8998
Qy 6733 AACTACGGCTACCTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAGCCAGTTACC 9058
Qy 6793 TTCCGAAAGAGGTTGGTAGCTCTTGATCCGGCAAAACAAACCCAGCTGGTAGCGGTGGT 6852
Db 9059 TTCCGAAAGAGGTTGGTAGCTCTTGATCCGGCAAAACAAACCCAGCTGGTAGCGGTGGT 9118
Qy 6853 TTTTGTGTTCCAAAGCAGCAGATTCAGCGGAGAAAAAGGATCTCAAGAGATCTCTTTG 6912
Db 9119 TTTTGTGTTCCAAAGCAGCAGATTCAGCGGAGAAAAAGGATCTCAAGAGATCTCTTTG 9178
Qy 6913 ATCTTTTCTACGGGCTGTGAGCTCAGTGGAAACAAACCTCAGCTTAAGGATTTTGGTC 6972
Db 9179 ATCTTTTCTACGGGCTGTGAGCTCAGTGGAAACAAACCTCAGCTTAAGGATTTTGGTC 9238
Qy 6973 ATGAGATTATCAAAAGGATCTTCACTAGATCTCTTTTAAATTAATAAGTTTTAAA 7032
Db 9239 ATGAGATTATCAAAAGGATCTTCACTAGATCTCTTTTAAATTAATAAGTTTTAAA 9298
Qy 7033 TCAATCTAAAGTATATATGAGTAACTTGTCTGACAGTTACCAATGCTTAATCAGTGAG 7092
Db 9299 TCAATCTAAAGTATATATGAGTAACTTGTCTGACAGTTACCAATGCTTAATCAGTGAG 9358
Qy 7093 GCACCTATCTCAGCGATCTGTCTATTTTCTATCTTCCATCTCCTGCTGCTGCTGCTG 7152
Db 9359 GCACCTATCTCAGCGATCTGTCTATTTTCTATCTTCCATCTCCTGCTGCTGCTGCTG 9418
Qy 7153 TAGATAACTAGATACGAGGAGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGA 7212
Db 9419 TAGATAACTAGATACGAGGAGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGA 9478
Qy 7213 GACCACTCTCAGCGCTCCAGATTTATCAGCAATTAACACAGCCAGCGGAGGCGGAG 7272
Db 9479 GACCACTCTCAGCGCTCCAGATTTATCAGCAATTAACACAGCCAGCGGAGGCGGAG 9538
Qy 7273 CGCAGAGTGTCTGCAACTTTATCCGCTCCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 7332
Db 9539 CGCAGAGTGTCTGCAACTTTATCCGCTCCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 9598
Qy 7333 GCTAGAGTAAAGTGTCTGCGGAGTAAAGTGTGCGCAAGTGTGCTGCTGCTGCTGCTGCTG 7392
Db 9599 GCTAGAGTAAAGTGTCTGCGGAGTAAAGTGTGCGCAAGTGTGCTGCTGCTGCTGCTGCTG 9658
Qy 7393 ATCTGTGTGTCAAGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7452
Db 9659 ATCTGTGTGTCAAGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9718
Qy 7453 AGCGAGATTAATGATCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7512
Db 9719 AGCGAGATTAATGATCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9778
Qy 7513 ATCTGTGTGTCAGAGTAAAGTGTGCGGAGTAAAGTGTGCGCAAGTGTGCTGCTGCTGCTGCTG 7572
Db 9779 ATCTGTGTGTCAGAGTAAAGTGTGCGGAGTAAAGTGTGCGCAAGTGTGCTGCTGCTGCTGCTG 9838
Qy 7573 AATCTCTTACTGTCTGATCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7632
Db 9839 AATCTCTTACTGTCTGATCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9898
Qy 7633 AAGTCAATCTGAGATAGTGTATGCGGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7692
Db 9899 AAGTCAATCTGAGATAGTGTATGCGGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9958
Qy 7693 GATAATACCGCCCATAGCAGAACTTTTAAAGTGTCTCATCTTGGAAAAAGTCTTCTG 7752
Db 9959 GATAATACCGCCCATAGCAGAACTTTTAAAGTGTCTCATCTTGGAAAAAGTCTTCTG 10018
Qy 7753 GGGCGAAATCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCACTGCT 7812
Db 10019 GGGCGAAATCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCACTGCT 10078
Qy 7813 GCACCCAACTGATCTTTCAGCATCTTTTACTTTTCCAGCGGTTTCTGGGTGAGCAAAAACA 7872
Db 10079 GCACCCAACTGATCTTTCAGCATCTTTTACTTTTCCAGCGGTTTCTGGGTGAGCAAAAACA 10138

Db 8999 AACTACGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGTGTAAGCCAGTTACC 9058
Qy 6793 TTCCGAAAGAGGTTGGTAGCTCTTGATCCGGCAAAACAAACCCAGCTGGTAGCGGTGGT 6852
Db 9059 TTCCGAAAGAGGTTGGTAGCTCTTGATCCGGCAAAACAAACCCAGCTGGTAGCGGTGGT 9118
Qy 6853 TTTTGTGTTCCAAAGCAGCAGATTCAGCGGAGAAAAAGGATCTCAAGAGATCTCTTTG 6912
Db 9119 TTTTGTGTTCCAAAGCAGCAGATTCAGCGGAGAAAAAGGATCTCAAGAGATCTCTTTG 9178
Qy 6913 ATCTTTTCTACGGGCTGTGAGCTCAGTGGAAACAAACCTCAGCTTAAGGATTTTGGTC 6972
Db 9179 ATCTTTTCTACGGGCTGTGAGCTCAGTGGAAACAAACCTCAGCTTAAGGATTTTGGTC 9238
Qy 6973 ATGAGATTATCAAAAGGATCTTCACTAGATCTCTTTTAAATTAATAAGTTTTAAA 7032
Db 9239 ATGAGATTATCAAAAGGATCTTCACTAGATCTCTTTTAAATTAATAAGTTTTAAA 9298
Qy 7033 TCAATCTAAAGTATATATGAGTAACTTGTCTGACAGTTACCAATGCTTAATCAGTGAG 7092
Db 9299 TCAATCTAAAGTATATATGAGTAACTTGTCTGACAGTTACCAATGCTTAATCAGTGAG 9358
Qy 7093 GCACCTATCTCAGCGATCTGTCTATTTTCTATCTTCCATCTCCTGCTGCTGCTGCTG 7152
Db 9359 GCACCTATCTCAGCGATCTGTCTATTTTCTATCTTCCATCTCCTGCTGCTGCTGCTG 9418
Qy 7153 TAGATAACTAGATACGAGGAGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGA 7212
Db 9419 TAGATAACTAGATACGAGGAGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGA 9478
Qy 7213 GACCACTCTCAGCGCTCCAGATTTATCAGCAATTAACACAGCCAGCGGAGGCGGAG 7272
Db 9479 GACCACTCTCAGCGCTCCAGATTTATCAGCAATTAACACAGCCAGCGGAGGCGGAG 9538
Qy 7273 CGCAGAGTGTCTGCAACTTTATCCGCTCCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 7332
Db 9539 CGCAGAGTGTCTGCAACTTTATCCGCTCCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 9598
Qy 7333 GCTAGAGTAAAGTGTCTGCGGAGTAAAGTGTGCGCAAGTGTGCTGCTGCTGCTGCTGCTG 7392
Db 9599 GCTAGAGTAAAGTGTCTGCGGAGTAAAGTGTGCGCAAGTGTGCTGCTGCTGCTGCTGCTG 9658
Qy 7393 ATCTGTGTGTCAAGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7452
Db 9659 ATCTGTGTGTCAAGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9718
Qy 7453 AGCGAGATTAATGATCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7512
Db 9719 AGCGAGATTAATGATCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9778
Qy 7513 ATCTGTGTGTCAGAGTAAAGTGTGCGGAGTAAAGTGTGCGCAAGTGTGCTGCTGCTGCTGCTG 7572
Db 9779 ATCTGTGTGTCAGAGTAAAGTGTGCGGAGTAAAGTGTGCGCAAGTGTGCTGCTGCTGCTGCTG 9838
Qy 7573 AATCTCTTACTGTCTGATCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7632
Db 9839 AATCTCTTACTGTCTGATCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9898
Qy 7633 AAGTCAATCTGAGATAGTGTATGCGGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7692
Db 9899 AAGTCAATCTGAGATAGTGTATGCGGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9958
Qy 7693 GATAATACCGCCCATAGCAGAACTTTTAAAGTGTCTCATCTTGGAAAAAGTCTTCTG 7752
Db 9959 GATAATACCGCCCATAGCAGAACTTTTAAAGTGTCTCATCTTGGAAAAAGTCTTCTG 10018
Qy 7753 GGGCGAAATCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCACTGCT 7812
Db 10019 GGGCGAAATCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCACTGCT 10078
Qy 7813 GCACCCAACTGATCTTTCAGCATCTTTTACTTTTCCAGCGGTTTCTGGGTGAGCAAAAACA 7872
Db 10079 GCACCCAACTGATCTTTCAGCATCTTTTACTTTTCCAGCGGTTTCTGGGTGAGCAAAAACA 10138

```

Qy 7873 GGAAGGCAAAATGCGGCAAAAGGAAATAGGCGGACACGAAATGTTGAATCTCAT 7932
Db 10139 GGAAGGCAAAATGCGGCAAAAGGAAATAGGCGGACACGAAATGTTGAATCTCAT 10198
Qy 7933 CTCCTCTCTTTTCAATATATTGAAGCATTTATCAGGGTTATCTCATGAGCGGATAC 7992
Db 10199 CTCCTCTCTTTTCAATATATTGAAGCATTTATCAGGGTTATCTCATGAGCGGATAC 10258
Qy 7993 ATATTGAAATGATTAGAAAAATAAACAATAGGGTTCCGCGCACATTTCCCGGAAAA 8052
Db 10259 ATATTGAAATGATTAGAAAAATAAACAATAGGGTTCCGCGCACATTTCCCGGAAAA 10318
Qy 8053 GTGCCACCTGAGCTTAAGAACCAATATTATCATGACATTAACCTATAAATAAGGCGT 8112
Db 10319 GTGCCACCTGAGCTTAAGAACCAATATTATCATGACATTAACCTATAAATAAGGCGT 10378
Qy 8113 ATCAGAGCGCCCTTCGCTCTTCAAGAAATTAATTTCTCATGTTTGACAGCTTATCATCGATA 8172
Db 10379 ATCAGAGCGCCCTTCGCTCTTCAAGAAATTAATTTCTCATGTTTGACAGCTTATCATCGATA 10438
Qy 8173 AGCTGACATCATGTTGGTATTGAAATAGAGCGAGATCGGAAACACTGAAAAATAACAGT 8232
Db 10439 AGCTGACATCATGTTGGTATTGAAATAGAGCGAGATCGGAAACACTGAAAAATAACAGT 10498
Qy 8233 TATTATTGG 8241
Db 10499 TATTATTGG 10507

```

```

RESULT 12
US-10-038-722-71
; Sequence 71, Application US/10038722
; Publication No. US20030175919A1
; GENERAL INFORMATION:
; APPLICANT: LEY, Arthur C.
; APPLICANT: GUTERMAN, Sonia K.
; APPLICANT: MARKLAND, William
; APPLICANT: KENT, Rachel B.
; APPLICANT: ROBERTS, Bruce L.
; APPLICANT: LADNER, Robert K.
; TITLE OF INVENTION: ITI-D1 KUNITZ DOMAIN MUTANTS AS nHE INHIBITORS
; FILE REFERENCE: LEV-1B
; CURRENT APPLICATION NUMBER: US/10/038,722
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 08/849,406
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: PCT/US95/16349
; PRIOR FILING DATE: 1995-12-15
; PRIOR APPLICATION NUMBER: US 08/358,160
; PRIOR FILING DATE: 1994-12-16
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 71
; LENGTH: 8584
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Plasmid pHIL-D2 (MPalpaPrePro::EPI-HNE-3) (Table 251)
US-10-038-722-71

```

```

Query Match 62.4%; Score 5142.8; DB 15; Length 8584;
Best Local Similarity 94.6%; Pred. No. 0;
Matches 5535; Conservative 0; Mismatches 167; Indels 152; Gaps 14;

Qy 2 GATCTAACATCCAAAGCAGAAAGGTTGAATGAAACCTTTTGGCCATCCGACATCCACAGG 61
Db 14 GATCTAACATCCAAAGCAGAAAGGTTGAATGAAACCTTTTGGCCATCCGACATCCACAGG 73
Qy 62 TCATTCTCACATTAAGTGCACAAACGCAACAGGAGGGGATACACTAGCAGCAGACCGTT 121
Db 74 TCATTCTCACATTAAGTGCACAAACGCAACAGGAGGGGATACACTAGCAGCAGACCGTT 133

```

```

Qy 122 GCAACGCGAGGACCTCCACTCTCTCTCTCAACACCCACCTTTTGGCCATCGAAAAACCA 181
Db 134 GCAACGCGAGGACCTCCACTCTCTCTCTCAACACCCACCTTTTGGCCATCGAAAAACCA 193
Qy 182 GCCCAGCTTATGGGCTTGATTGAGAGCTCGCTCATTTCCAAATTCCTTCTATTAGGCTACTAA 241
Db 194 GCCCAGCTTATGGGCTTGATTGAGAGCTCGCTCATTTCCAAATTCCTTCTATTAGGCTACTAA 253
Qy 242 CACCATGACTTTATAGCTGTCTATCCCTGCGCCCTCGCGAGGTTCTATGTTGTTTAT 301
Db 254 CACCATGACTTTATAGCTGTCTATCCCTGCGCCCTCGCGAGG-TCAITGTTGTTTAT 312
Qy 302 TTCGGAATGCAACAGCTCCGATTAACACCCGAAACATCACTCCAGATGAGGGCTTTCGA 361
Db 313 TTCGGAATGCAACAGCTCCGATTAACACCCGAAACATCACTCCAGATGAGGGCTTTCGA 372
Qy 362 GTGTGGGCTCAATATGTTTCATGTTCCCAATAGGCCCAAAACTGACAGTTTAAAGCGTG 421
Db 373 GTGTGGGCTCAATATGTTTCATGTT-CCCAATAGGCCCAAAACTGACAGTTTAAAGCGTG 431
Qy 422 TCTTGGAACTTAATAAGCAAAAGCGTGATCTCATCCAAAGATCAACTAAAGTTTGGTTCGT 481
Db 432 TCTTGGAACTTAATAAGCAAAAGCGTGATCTCATCCAAAGATCAACTAAAGTTTGGTTCGT 491
Qy 482 TGAATGCTAACCGCCAGTTGGTCAAAAAGAAATTTCCAAAAGTCCGCCATACCGTTTGTG 541
Db 492 TGAATGCTAACCGCCAGTTGGTCAAAAAGAAACTTCCAAAAGTCCGCCATACCGTTTGTG 551
Qy 542 TTGTTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAATGCTTAGCGCAGTCTC 601
Db 552 TTGTTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAATGCTTAGCGCAGTCTC 611
Qy 602 TCTATCGCTTCTGAAACCCCGGTGCACCTGTGCGGAAACGCAAAATGGGGAAC-ACCCGCT 660
Db 612 TCTATCGCTTCTGAAACCCCGGTGCACCTGTGCGGAAACGCAAAATGGGGAACACACCGCT 671
Qy 661 TTTTGGATGATTATGATTTGT-CTCACAATTTGATGCTTCCAAAGATCTGGTGGGATAC 719
Db 672 TTTTGGATGATTATGATTTGTCTCCACATTTGATGCTTCCAAAGATCTGGTGGGATAC 731
Qy 720 TGCTGATAGCTTAACGTTTCAATGATCAAAATTTAACTGTTCTTAACCCCTACTTGACAC- 778
Db 732 TGCTGATAGCTTAACGTTTCAATGATCAAAATTTAACTGTTCTTAACCCCTACTTGACAGGCA 791
Qy 779 ATATATAACAGAGGAGCTGCCCTGTCTTAAACCTTTTATTTTATCATATTATAGC 838
Db 792 ATATATAACAGAGGAGCTGCCCTGTCTTAAACCTTTTATTTATCATATTATAGC 851
Qy 839 TTACTTTTCAATAATTGGGACTGTTTCCAAATTTGACAAAGCTTTTGAATTTTAAAGCTTTTAA 898
Db 852 TTACTTTTCAATAATTGGGACTGTTTCCAAATTTGACAAAGCTTTTGAATTTTAAAGCTTTTAA 911
Qy 899 GACAACTTGAGAGATCAAAAAACAACTAATTTTGAAGGATCCAAACGATGAGATTTC 958
Db 912 GACAACTTGAGAGATCAAAAAACAACTAATTTTTCG-----AAACGATGAGATTTC 963
Qy 959 CTTCAATTTTACTGCAATTTTATTTGCGAGCATCTCCGCAATAGCTGCTCCAGTCAACA 1018
Db 964 CATCTATCTTCACTGCTGTTTGTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1023
Qy 1019 CTACAACAGAGATGAAACCGGCAAAATTCGGCTGAAGCTGTCATCGGTTACTTCAGATT 1078
Db 1024 CCCTACTTGAGAGAGAGAGCTGCTCAAAATCTCTGCTGAGGCTGCTGCTGCTGCTGCTGCTG 1083
Qy 1079 TAGAAGGGGATTTGATGTTGCTGTTTGGCCATTTTCCAAACAGCACAATAAAGCGGTTAT 1138
Db 1084 TGAAGGCTGACTTCGAGCTGCTGTTTGGCCATTTTCTAACTCTACTTAAACAGCGTTTGT 1143
Qy 1139 TGTTTATAAATCTACTATTGCGAGCATCTGCTAAAGAGAGAGGGTATCTCTCCAGA 1198
Db 1144 TGTTCATCAACACTTACCATCGCTTCTATCGCTGCTAAGGAGGAGGTTGTTTCTTGGACA 1203
Qy 1199 AAGAGAGGCTGAGAGCCCGAGGCCCGAGGCCCGAGGCCCGAGGCCCGAGGAAACATGTATCA 1258

```


QY	3419	CGTTTTTGGATAGACGACGAGGCGATCTTCAATTTCTTGTGAGGAGGCTTAGAAC	3478
Db	3348	CGTTTTTGGATAGACGACGAGGCGATCTTCAATTTCTTGTGAGGAGGCTTAGAAC	3407
QY	3479	GTCAATTTTGCACAAATCAATACGACCTTCAGAGGAGCTTCTTTAGGTTTGGATTTTC	3538
Db	3408	GTCAATTTTGCACAAATCAATACGACCTTCAGAGGAGCTTCTTTAGGTTTGGATTTTC	3467
QY	3539	TTTAGGTTTGGTCTTGGTGTATCTTGGCTTGGCATCTCTTCTTCTAGTGACCTTTAG	3598
Db	3468	TTTAGGTTTGGTCTTGGTGTATCTTGGCTTGGCATCTCTTCTTCTAGTGACCTTTAG	3527
QY	3599	GGACTTCATATCCAGGTTTCTCTCCACCTCGTCCAAAGCTCACCGTACTTTGGCACATCT	3658
Db	3528	GGACTTCATATCCAGGTTTCTCTCCACCTCGTCCAAAGCTCACCGTACTTTGGCACATCT	3587
QY	3659	AACTAATGCAAAATAAATAAGTACGACATTTCCAGGCTATACTTCTTGGATTTAGC	3718
Db	3588	AACTAATGCAAAATAAATAAGTACGACATTTCCAGGCTATACTTCTTGGATTTAGC	3647
QY	3719	TTCTCCAGTTTCATCAGCTTCTCTCTTAAATTTTAGGTTTCAACAAATCTTGGTCTCAAA	3778
Db	3648	TTCTCCAGTTTCATCAGCTTCTCTCTTAAATTTTAGGTTTCAACAAATCTTGGTCTCAAA	3707
QY	3779	TAAACGTTTGGTATAGAAACCTTCTGGAGCATTTGCTTTAGCATCCCAAGGTGGCTTC	3838
Db	3708	TAAACGTTTGGTATAGAAACCTTCTGGAGCATTTGCTTTAGCATCCCAAGGT - GCTTC	3766
QY	3839	GATGGCTCTAAGACCTTTGATGGCCAAAGCAGAGTGGTTCCTAAGTACGACGAAACC	3898
Db	3767	CATGGCTCTAAGACCTTTGATGGCCAAAGCAGAGTGGTTCCTAAGTACGACGAAACC	3826
QY	3899	AAACCTGTTTGTCTCAACACAAATTTCAAGCAGTCTCATCACAAATCCAAATTCGATACC	3958
Db	3827	AAACCTGTTTGTCTCAACACAAATTTCAAGCAGTCTCATCACAAATCCAAATTCGATACC	3886
QY	3959	CAGCAACTTTTGGTGTCTCAGATGTAGCACTTTATACCAAAACCGTACGACGAGA	4018
Db	3887	CAGCAACTTTTGGTGTCTCAGATGTAGCACTTTATACCAAAACCGTACGACGAGA	3946
QY	4019	TTGTGTAGCTCCAGTTTGTCTTATAGCTCCGGAATAGACTTTTGGAGAGTACAC	4078
Db	3947	TTGTGTAGCTCCAGTTTGTCTTATAGCTCCGGAATAGACTTTTGGAGAGTACAC	4006
QY	4079	CAGGCCCAACGATTAATAGAGAGTACGCCAACCAAGTGTGAATAGACCATCGGGCG	4138
Db	4007	CAGGCCCAACGATTAATAGAGAGTACGCCAACCAAGTGTGAATAGACCATCGGGCG	4066
QY	4139	GTCAAGTACCAAGACGCCCAACAAATTTTCACTGACAGGAATTTTGGACATCTCAGA	4198
Db	4067	GTCAAGTACCAAGACGCCCAACAAATTTTCACTGACAGGAATTTTGGACATCTCAGA	4126
QY	4199	AACTTCGTATTCAGTGTCAATTTCCGAGCATCAATAATGGGGATTTATACCAAGAGCAAC	4258
Db	4127	AACTTCGTATTCAGTGTCAATTTCCGAGCATCAATAATGGGGATTTATACCAAGAGCAAC	4186
QY	4259	AGTGGAGTACATCTACCAACTTTTGGCTCTCAGAAAGGATTAACAGTTCTACTACC	4318
Db	4187	AGTGGAGTACATCTACCAACTTTTGGCTCTCAGAAAGGATTAACAGTTCTACTACC	4246
QY	4319	GCCATTAGTGAATTTTCAATTCGCCAGTGGAGAGAAAGGACACAGCATACTAGC	4378
Db	4247	GCCATTAGTGAATTTTCAATTCGCCAGTGGAGAGAAAGGACACAGCATACTAGC	4306
QY	4379	ATTAGCGGCAAGGATGCAACTTTTATCAACCGGCTCTATAGATTAACCTAGCGCTGG	4438
Db	4307	ATTAGCGGCAAGGATGCAACTTTTATCAACCGGCTCTATAGATTAACCTAGCGCTGG	4366
QY	4439	GATCATCTTTGACAACTTTTCTGCAATCTAGTCCAAATCTCAATCTGATACC	4498
Db	4367	GATCATCTTTGACAACTTTTCTGCAATCTAGTCCAAATCTCAATCTGATACC	4426

QY	4499	ATTATTTGTACAACTTGAGCAAGTTGTGATCAGCTCCTCAAAATTGGTCTCTGTACCGGA	4558
Db	4427	ATTATTTGTACAACTTGAGCAAGTTGTGATCAGCTCCTCAAAATTGGTCTCTGTACCGGA	4436
QY	4559	TGACTCAACTTGCACATTAACCTTGAAGCTCAGTCGATGAGTGAACTTGATCAGGTTGTG	4618
Db	4437	TGACTCAACTTGCACATTAACCTTGAAGCTCAGTCGATGAGTGAACTTGATCAGGTTGTG	4496
QY	4619	CAGTGTGAGCAGCAGTAGGAAACACCGCTTTTCTACCAACTCAAGGAAATATCAAA	4678
Db	4497	CAGTGTGAGCAGCAGTAGGAAACACCGCTTTTCTACCAACTCAAGGAAATATCAAA	4556
QY	4679	CTCTGAAACACTTGCCTGATAGCAGGTAGCAAGGAAATGTCTACTTGAAGTCCGACAGTG	4738
Db	4557	CTCTGAAACACTTGCCTGATGAGGTAGCAAGGAAATGTCTACTTGAAGTCCGACAGTG	4616
QY	4739	AGTGTAGTCTTGAAGAAATCTGAAGCCGTATTTTATTAATCAGTGAGTCAATCAGGA	4798
Db	4617	AGTGTAGTCTTGAAGAAATCTGAAGCCGTATTTTATTAATCAGTGAGTCAATCAGGA	4676
QY	4799	GATCCTCTACGCCGCGACGATCGTGGCCGACCTGCGAGTCTGGCATCACCGGCGCCACAGG	4858
Db	4677	GATCCTCTACGCCGCGACGATCGTGGC-----CGGCATCACCGGCGCCACAGG	4724
QY	4859	TGGGTTGCTGGGCCCTATATATCGCCGACATCAACGATGGGGAAGATCGGGCTCGCCACTT	4918
Db	4725	TGGGTTGCTGGGCCCTATATATCGCCGACATCAACGATGGGGAAGATCGGGCTCGCCACTT	4784
QY	4919	CGGCTCATGAGCGCTTGTTCGGCGTGGGTATGTGGCAGGCGCCGCTGGCGGGGACT	4978
Db	4785	CGGCTCATGAGCGCTTGTTCGGCGTGGGTATGTGGCAGGCGCCGCTGGCGGGGACT	4844
QY	4979	GTTGGGGCGCATCTCTTGTGATGACCATTCCTTGGCGGGCGGCTGCTCAACGGCCTCAA	5038
Db	4845	GTTGGGGCGCATCTCTTGTGATGACCATTCCTTGGCGGGCGGCTGCTCAACGGCCTCAA	4904
QY	5039	CCTACTACTGGCTGCTTCTTAATGAGGAGTGCCTAAGCGAGAGCGTCCAGTATCTAT	5098
Db	4905	CCTACTACTGGCTGCTTCTTAATGAGGAGTGCCTAAGCGAGAGCGTCCAGTATCTAT	4964
QY	5099	GATTGGAAGTATGGGAAATGGTGATACCCGCAATCTTCAGTGTCTTGAGGTCCTCTATCAG	5158
Db	4965	GATTGGAAGTATGGGAAATGGTGATACCCGCAATCTTCAGTGTCTTGAGGTCCTCTATCAG	5024
QY	5159	ATTATGCCCAACTTAAGCAACCGGAGGAGGAGATTTTCATGTAATTTCTCTGACTTTTG	5218
Db	5025	ATTATGCCCAACTTAAGCAACCGGAGGAGGAGATTTTCATGTAATTTCTCTGACTTTTG	5084
QY	5219	GTCACTCAGTAGACTCGAACTGTGAGACTATCTCGGTTTATGACAGCAGAAATGCTCTTCTT	5278
Db	5085	GTCACTCAGTAGACTCGAACTGTGAGACTATCTCGGTTTATGACAGCAGAAATGCTCTTCTT	5144
QY	5279	GGAGACAGTAAATGAAGTCCCAACCAATAAAGAAATCCTTGTATCAGGAACAACTTCTT	5338
Db	5145	GGAGACAGTAAATGAAGTCCCAACCAATAAAGAAATCCTTGTATCAGGAACAACTTCTT	5204
QY	5339	GTTTCGAACTTTTTCGGTGCTTGAATATAAATGATAGAGTGGATATGTCGGGTAGGAA	5398
Db	5205	GTTTCGAACTTTTTCGGTGCTTGAATATAAATGATAGAGTGGATATGTCGGGTAGGAA	5264
QY	5399	TGGAGCGGCAAAATGCTTACTTCTGCACTTCAAGAGGTATGATAGGTTTGTAGTACT	5458
Db	5265	TGGAGCGGCAAAATGCTTACTTCTGCACTTCAAGAGGTATGATAGGTTTGTAGTACT	5324
QY	5459	GATGCCAACTTCAGTGACAGCTTGTATTTTCGTTCAAAACCTTCCGAATCCAGAGAAAT	5518
Db	5325	GATGCCAACTTCAGTGACAGCTTGTATTTTCGTTCAAAACCTTCCGAATCCAGAGAAAT	5384
QY	5519	CAAAAGTTTGTCTACTTACTTATTCAGGCGAGTGGCTCTTGAAATCTGAAATAGTGTG	5578
Db	5385	CAAAAGTTTGTCTACTTACTTATTCAGGCGAGTGGCTCTTGAAATCTGAAATAGTGTG	5444
QY	5579	CTCGTTTGTGAGTGTATCTTTGTATGAATAAATCTAGTCTTTCATCTAAATATCTTGA	5638

5445	Db	CTCGTGTGTTTGAGGTCACTCTTTGTATGAATAAATCTAGTCTTTTGATCTAAATAATCTTGA	5504
5439	Qy	CGAGCCAAGGCGATAAATACCCAAATCTAAAACTCTTTTAAAAAGTCTTAAAAAGGCAAGTA	5698
5505	Db	CGAGCCAAGGCGATAAATACCCAAATCTAAAACTCTTTTAAAAAGTCTTAAAAAGGCAAGTA	5564
5639	Qy	TGTCGTGCTGTATTAACCCCAAAATCAGCTCGTAGTCTGATCCTCATCAACTTGAAGGGC	5758
5565	Db	TGTCGTGCTGTATTAACCCCAAAATCAGCTCGTAGTCTGATCCTCATCAACTTGAAGGGC	5624
5759	Qy	ACTATCTTGTTTTAGAGAAATTTGCGGAGATCGGATATCGAGAAAAAGGTACGCTGATTT	5818
5625	Db	ACTATCTTGTTTTAGAGAAATTTGCGGAGATCGGATATCGAGAAAAAGGTACGCTGATTT	5684
5819	Qy	TAAACGTGAAATTTTATCTCAAGATCTCTGCCTCG	5852
5685	Db	TAAACGTGAAATTTTATCTCAAGATCTCGGCGCGG	5718

```

RESULT 13
US-10-038-722-75
; Sequence 75, Application US/10038722
; Publication No. US20030175919A1
; GENERAL INFORMATION:
; APPLICANT: LEY, Arthur C.
; APPLICANT: GUTERMAN, Sonia K.
; APPLICANT: MARKLAND, William
; APPLICANT: KENT, Rachel B.
; APPLICANT: ROBERTS, Bruce L.
; APPLICANT: LADNER, Robert C.
; TITLE OF INVENTION: ITI-D1 KUNITZ DOMAIN MUTANTS AS DHE INHIBITORS
; FILE REFERENCE: LEY=13
; CURRENT APPLICATION NUMBER: US/10/038,722
; CURRENT FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 08/849,406
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: PCT/US95/16349
; PRIOR FILING DATE: 1995-12-15
; PRIOR APPLICATION NUMBER: US 08/358,160
; PRIOR FILING DATE: 1994-12-16
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75
; LENGTH: 8590
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pd2pick (MfalpaaPrePro::BPI-NHE-3) circular ds
US-10-038-722-75

```

[illegible]

1379 CCATCGAGAACACTCAAGAGAGAAATGTCCCTTCTAACTAGTGGCGTAGAAATCCCTAG 1438
1313 TTGTCAGAGTAAC---GGTAACAAGTCTACTCTGAGAAGGAGTGTAGAGAGTACTGTGG 1369
1439 GCGCGCGGGAATTAATTCGCTCTAGACATGACTGTTCTCTAGTTCAGTTCAGTTCGGCACTTA 1498
1370 TGTTCATATAGAAATTCGCTCTAGACATGACTGTTCTCTAGTTCAGTTCAGTTCGGCA-TTA 1428
1499 CGAAGAACCGGTCCTTCTAGATTTCTAATCAAGAGAGATGTCAAGATGCCATTTGGCTGAG 1558
1429 CGAAGAACCGGTCCTTCTAGATTTCTAATCAAGAGAGATGTCAAGATGCCATTTGGCTGAG 1488
1559 AGATGAGGCTTCATTTTGTATCTTTTATTTTATTTGTAACCTATATATAGTATAGATTTT 1618
1489 AGATGAGGCTTCATTTTGTATCTTTTATTTTATTTGTAACCTATATATAGTATAGATTTT 1548
1619 TTGTCATTTTGTCTTCTCTGACGAGCTTGCTCTGATCAGCTATCTCGCAGCTGATG 1678
1549 TTGTCATTTTGTCTTCTCTGACGAGCTTGCTCTGATCAGCTATCTCGCAGCTGATG 1608
1679 AATATCTTGTGTAGGGTTTGGGAAATCATTCGAGTTTGATGTTTCTTGGTATTTTC 1738
1609 AATATCTTGTGTAGGGTTTGGGAAATCATTCGAGTTTGATGTTTCTTGGTATTTTC 1668
1739 CCACTCTCTCAGAGTACAGAGATTAAGTGAGAGATTCGTTTGTGCAAGCTTATCGAT 1798
1669 CCACTCTCTCAGAGTACAGAGATTAAGTGAGAGATTCGTTTGTGCAAGCTTATCGAT 1728
1799 AAGCTTTAATGCGGTAGTTTATCAGAGTTAAATTCGTAACGAGTCAGGACCGGTATG 1858
1729 AAGCTTTAATGCGGTAGTTTATCAGAGTTAAATTCGTAACGAGTCAGGACCGGTATG 1788
1859 AATCTAAAGTCCGCTATCGTATCTCTCGGACCGTACCGCTGATGCTAGGCATA 1918
1789 AATCTAAAGTCCGCTATCGTATCTCTCGGACCGTACCGCTGATGCTAGGCATA 1848
1919 GCGTTGGTTATGCGGTACTGCGGGGCTCTTGGCGGATATCGTCCATTCGACAGCATC 1978
1849 GCGTTGGTTATGCGGTACTGCGGGGCTCTTGGCGGATATCGTCCATTCGACAGCATC 1908
1979 GCGAGTCACTATGCGGTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCC 2038
1909 GCGAGTCACTATGCGGTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCC 1968
2039 GTTCTCGGAGCATGTCGACCGCTTTGGCGCGCCGACGCTCTGCTCGCTTCTGCTACTT 2098
1969 GTTCTCGGAGCATGTCGACCGCTTTGGCGCGCCGACGCTCTGCTCGCTTCTGCTACTT 2028
2099 GGAGCCACTATCGACTACCGCATATGCGGACCAACCGGTCCTGTGGATCTATCGAATC 2158
2029 GGAGCCACTATCGACTACCGCATATGCGGACCAACCGGTCCTGTGGATCTATCGAATC 2088
2159 TAAATGTAAGTTAAATCTCTAAATTAATTAATTAAGTCCAGTTTCTCCATACGAACCTT 2218
2089 TAAATGTAAGTTAAATCTCTAAATTAATTAATTAAGTCCAGTTTCTCCATACGAACCTT 2148
2219 AACAGCATTCGCGTGAGCATCTAGACCTTCAACAGCAGCCAGATCCATCACTGTGGCC 2278
2149 AACAGCATTCGCGTGAGCATCTAGACCTTCAACAGCAGCCAGATCCATCACTGTGGCC 2208
2279 AATATGTTTTCAGTCCCTCAGGAGTTACGCTTGTGAAGTATGAACTCTTGGAGGTTGC 2338
2209 AATATGTTTTCAGTCCCTCAGGAGTTACGCTTGTGAAGTATGAACTCTTGGAGGTTGC 2268
2339 AGTGTTAACCTCGCTGATTTAGCGGCATATCCGTACGTTTGGCAAGTGTGGTTGCTACC 2398
2269 AGTGTTAACCTCGCTGATTTAGCGGCATATCCGTACGTTTGGCAAGTGTGGTTGCTACC 2328
2399 GGAGGAGTAAATCTCCAACTCTCTGGAGAGTAGGCACCAACAAACACAGATCCAGCGTG 2458
2329 GGAGGAGTAAATCTCCAACTCTCTGGAGAGTAGGCACCAACAAACACAGATCCAGCGTG 2388

2459 TTGTACTTGATCAACATAAGAAGAAGCATTTCTCGATTTGCGAGGATCAAGTGTTCAGGAGC 2518
2389 TTGTACTTGATCAACATAAGAAGAAGCATTTCTCGATTTGCGAGGATCAAGTGTTCAGGAGC 2448
2519 GTACTGATTCGACATTTCCAAAGCCTGCTGAGTGTGCAACCGATAGGTTGTAGAGTG 2578
2449 GTACTGATTCGACATTTCCAAAGCCTGCTGAGTGTGCAACCGATAGGTTGTAGAGTG 2508
2579 TSCAATACACTTGGGTACAAATTTCAAACCTTGGCAACTGACACAGCTTGGTTGTGAACAGC 2638
2509 TSCAATACACTTGGGTACAAATTTCAAACCTTGGCAACTGACACAGCTTGGTTGTGAACAGC 2568
2639 ATCTTCAATTTCTGCAAGCTCTTGTCTGTCATATCGACAGCCAAAGATCACTCTGGGA 2698
2569 ATCTTCAATTTCTGCAAGCTCTTGTCTGTCATATCGACAGCCAAAGATCACTCTGGGA 2628
2699 ATCAATACACTTGGGTACAAATTTCAAACCTTGGCAACTGACACAGCTTGGTTGTGAACAGC 2758
2629 ATCAATACACTTGGGTACAAATTTCAAACCTTGGCAACTGACACAGCTTGGTTGTGAACAGC 2687
2759 TTTATCGCAATTAATACTAGAACTTTCAGAAAGCCGACAGGCAATGTCAATATACACAGGC 2818
2688 TTTATCGCAATTAATACTAGAACTTTCAGAAAGCCGACAGGCAATGTCAATATACACAGGC 2747
2819 TGATGTGTCAATTTTGAACCATCATCTTGGCAGCAGTAACGAACTGTGTTCTCTGACCAAA 2878
2748 TGATGTGTCAATTTTGAACCATCATCTTGGCAGCAGTAACGAACTGTGTTCTCTGACCAAA 2807
2879 TATTTGTTCACACTTAGGAAACAGTTTCTGTTCCGTAGCCATAGCAGCTACTGCTGGGC 2938
2808 TATTTGTTCACACTTAGGAAACAGTTTCTGTTCCGTAGCCATAGCAGCTACTGCTGGGC 2867
2938 GCCTTCTGCTAGCAGCATACACTTAGCACCACCTTGTGGGCAACGTAGATGATCTTCTGG 2998
2868 GCCTTCTGCTAGCAGCATACACTTAGCACCACCTTGTGGGCAACGTAGATGATCTTCTGG 2927
2999 GGTAAAGGTACCATCTTCTTAGTGGAGATGCAAAAACAAATTTCTTGTGAACAGCAAC 3058
2928 GGTAAAGGTACCATCTTCTTAGTGGAGATGCAAAAACAAATTTCTTGTGAACAGCAAC 2987
3059 TTTGGCAGGAAACACCCAGCATCAGGAAAGTGGAAAGCAGAAATTTGCGGTTCCACAGGAAT 3118
2988 TTTGGCAGGAAACACCCAGCATCAGGAAAGTGGAAAGCAGAAATTTGCGGTTCCACAGGAAT 3047
3119 ATAGAGGCCAACTTTCTCAATAGTCTTGAACAGCAGACAGCATACACAGGCGAAGT 3178
3048 ATAGAGGCCAACTTTCTCAATAGTCTTGAACAGCAGACAGCATACACAGGCGAAGT 3107
3179 CTCAACTTGAACAGCTCTCCGTTAGTTGAGCTTTCATGGAATTTCTGACGTTATCTATAGA 3238
3108 CTCAACTTGAACAGCTCTCCGTTAGTTGAGCTTTCATGGAATTTCTGACGTTATCTATAGA 3167
3239 GAGATCAATGGCTCTCTTAACGTTATCTGCAATTTGCAATAGTTCCTCTGGGAAGAGC 3298
3168 GAGATCAATGGCTCTCTTAACGTTATCTGCAATTTGCAATAGTTCCTCTGGGAAGAGC 3227
3299 TTCTAAACAGAGGTCTCTTCAAGCGACTCCATCAAACTTGGCAGTTAGTTCTAAAAGGCG 3358
3228 TTCTAAACAGAGGTCTCTTCAAGCGACTCCATCAAACTTGGCAGTTAGTTCTAAAAGGCG 3287
3359 TTTGTACCAATTTTGACGAAATTTGTCGAAATTTGTTTGAATTAATTCGATATCTGTTTC 3418
3288 TTTGTACCAATTTTGACGAAATTTGTCGAAATTTGTTTGAATTAATTCGATATCTGTTTC 3347
3419 CGTTTCTGATAGGAGCAGGAGGCTCTTCAATTTCTTGTGAGAGGCTTTAGAAAC 3478
3348 CGTTTCTGATAGGAGCAGGAGGCTCTTCAATTTCTTGTGAGAGGCTTTAGAAAC 3407
3479 GTCAATTTTGCACAAATTCATACGACTTCAGAAAGGAGCTTTCTTATAGTTTGAATCTTC 3538
3408 GTCAATTTTGCACAAATTCATACGACTTCAGAAAGGAGCTTTCTTATAGTTTGAATCTTC 3467
3539 TTTAGGTTGTTCTTGGTGTATCTCTGGCTTGGCAATCTCTTTCTTCTTAGTGACCTTAG 3598

Db 3468 TTTAGTTGTTCTTGGTATCTGGCTGGGATCTCTCTTCTTCTAGTACCTTTAG 3527
Qy 3599 GGACTTCATATCAGGTTTCTTCCACCTCGTCCAAAGTGCACACCGTACTTGGGCACTC 3658
Db 3528 GGACTTCATATCAGGTTTCTTCCACCTCGTCCAAAGTGCACACCGTACTTGGGCACTC 3587
Qy 3659 AACTAAATGCAAAATAAATAAGTCAGCACATCCAGGCTATATCTTCTTGGATTTAGC 3718
Db 3598 AACTAAATGCAAAATAAATAAGTCAGCACATCCAGGCTATATCTTCTTGGATTTAGC 3647
Qy 3719 TTCTGGAAGTTATCAGCTTCTTCCCTTAATTTAGGCTTCAACAAATCTTCTGCTCAAA 3778
Db 3648 TTCTGGAAGTTATCAGCTTCTTCCCTTAATTTAGGCTTCAACAAATCTTCTGCTCAAA 3707
Qy 3779 TAACCGTTTGGTATAAGAACCTTCTGGAGCATTCCTTACGATCCCAAGGTCGCTTC 3838
Db 3708 TAACCGTTTGGTATAAGAACCTTCTGGAGCATTCCTTACGATCCCAAGGTCGCTTC 3766
Qy 3839 CATGGCTCTAAGACCTTTGATTTGGCCAAAACAGGAAGTGGCTTCAAGTGCAGAAACC 3898
Db 3767 CATGGCTCTAAGACCTTTGATTTGGCCAAAACAGGAAGTGGCTTCAAGTGCAGAAACC 3826
Qy 3899 AACACCTGTTTGTTCACCAACAAATTTCAAGCAGTCTCCATCACAATCCAAATTCGATACC 3958
Db 3827 AACACCTGTTTGTTCACCAACAAATTTCAAGCAGTCTCCATCACAATCCAAATTCGATACC 3886
Qy 3959 CAGCAACTTTTGTGCTCCAGATGAGCACTTTTATACCAAAACCGTGCAGCAGAGA 4018
Db 3887 CAGCAACTTTTGTGCTCCAGATGAGCACTTTTATACCAAAACCGTGCAGCAGAGA 3946
Qy 4019 TTGCTAGACTCCAGTTTGTGCTTATAGCTCCGGAATAGACTTTTGGAGGAGTACAC 4078
Db 3947 TTGCTAGACTCCAGTTTGTGCTTATAGCTCCGGAATAGACTTTTGGAGGAGTACAC 4006
Qy 4079 CAGGCCAAACGAGTAATTAAGAGTGCAGCCACCAAGTGTGAATAGACCATCGGGGCG 4138
Db 4007 CAGGCCAAACGAGTAATTAAGAGTGCAGCCACCAAGTGTGAATAGACCATCGGGGCG 4066
Qy 4139 GTCAAGTGTCAAGAGCCCAACAAATTTCACTGCAGGGAATCTTTTGAATCTTCAGA 4198
Db 4067 GTCAAGTGTCAAGAGCCCAACAAATTTCACTGCAGGGAATCTTTTGAATCTTCAGA 4126
Qy 4199 AAGTTCGTATTCAGTGTCAATTTCCGAGCATCAATTAATGGGATTTATACAGAGCAAC 4258
Db 4127 AAGTTCGTATTCAGTGTCAATTTCCGAGCATCAATTAATGGGATTTATACAGAGCAAC 4186
Qy 4259 AGTGGAGTCAACATCTACCAACTTTTCCGCTCTCAGAAAAGCATAAAGTCTTACTACC 4318
Db 4187 AGTGGAGTCAACATCTACCAACTTTTCCGCTCTCAGAAAAGCATAAAGTCTTACTACC 4246
Qy 4319 GCCATTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAAGGCAAGGATACTAGC 4378
Db 4247 GCCATTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAAGGCAAGGATACTAGC 4306
Qy 4379 ATTAGCGGGCAGGATGCAACTTTTATACACAGGCTCTATAGATAACCTTAGCGCTGG 4438
Db 4307 ATTAGCGGGCAGGATGCAACTTTTATACACAGGCTCTATAGATAACCTTAGCGCTGG 4366
Qy 4439 GATCATCTTTTGGCAACCTTTTCTGCCAAATCTTAGGTCGCAAAATCACTTCATTGATACC 4498
Db 4367 GATCATCTTTTGGCAACCTTTTCTGCCAAATCTTAGGTCGCAAAATCACTTCATTGATACC 4426
Qy 4499 ATTATTGTACAACTTTCAGCAAGTTTGTGATCAGTCTCCTCAATTTGGTCTCTGTAAACGGA 4558
Db 4427 ATTAT-----ACGGA 4436
Qy 4559 TGACTCAACTTGCACTTAATTAAGTGCAGTGCAGTGAAGTGAAGTGAAGTGAAGTGAAG 4618
Db 4437 TGACTCAACTTGCACTTAATTAAGTGCAGTGCAGTGAAGTGAAGTGAAGTGAAGTGAAG 4496
Qy 4619 CAGCTGGTCAGGAGCATAGGAAAACAGGCTTTTCTTACCAAACTCAAGGAATTTATCAA 4678

Db 4497 CAGCTGGTCAGGAGCATAGGAAAACAGGCTTTTCTTACCAAACTCAAGGAATTTATCAA 4556
Qy 4679 CTCCTGCAACACTTGGGTATGACAGGTAGCAAGGAAAATCTCATATCTTGAAGTTCGACAGTG 4738
Db 4557 CTCCTGCAACACTTGGGTATGACAGGTAGCAAGGAAAATCTCATATCTTGAAGTTCGACAGTG 4616
Qy 4739 AGTGTAGTCTTGAGAAATTTCTGAAGCCGTATTTTATTTATCAGTGAAGTCACTCATCAGGA 4798
Db 4617 AGTGTAGTCTTGAGAAATTTCTGAAGCCGTATTTTATTTATCAGTGAAGTCACTCATCAGGA 4676
Qy 4799 GATCTCTTACGCGGAGCATCTGTCGCGCACTCTGCGAGTCTGGCATCAACCGGCGGCAACAGG 4858
Db 4677 GATCTCTTACGCGGAGCATCTGTCGCG-----CGGCACTCAACCGGCGGCAACAGG 4724
Qy 4859 TGGGTTGCTGGCGCTATATCGCGACATCAACGATGGGAAGATCGGGCTCGCCACTT 4918
Db 4725 TGGGTTGCTGGCGCTATATCGCGACATCAACGATGGGAAGATCGGGCTCGCCACTT 4784
Qy 4919 CGGGCTCATAGAGCGCTTGTTCGCGTGGGTATGCTGCGAGGCCCGCTGGCGGGGACT 4978
Db 4785 CGGGCTCATAGAGCGCTTGTTCGCGTGGGTATGCTGCGAGGCCCGCTGGCGGGGACT 4844
Qy 4979 GTTGGCGGCACTCTCTTGTGATGACCACTTCTTGGCGCGCGGTGTCTCAACGGCTCAA 5038
Db 4845 GTTGGCGGCACTCTCTTGTGATGACCACTTCTTGGCGCGCGGTGTCTCAACGGCTCAA 4904
Qy 5039 CCTACTCTGGGCTGCTTCTTAATGCGAGATGCGGATAAGGAGAGCGTGAAGTATCTAT 5098
Db 4905 CCTACTCTGGGCTGCTTCTTAATGCGAGATGCGGATAAGGAGAGCGTGAAGTATCTAT 4964
Qy 5099 GATTGGAAGTATGGAAATGATGATACCGCATCTCTTCAAGTGTCTTGAAGTCTCTCATCAG 5158
Db 4965 GATTGGAAGTATGGAAATGATGATACCGCATCTCTTCAAGTGTCTTGAAGTCTCTCATCAG 5024
Qy 5159 ATTATGCCCACTAAAGCAACCGGAGGAGAGATTTTCAATGTAATTTCTTCTGACTTTTG 5218
Db 5025 ATTATGCCCACTAAAGCAACCGGAGGAGAGATTTTCAATGTAATTTCTTCTGACTTTTG 5084
Qy 5219 GTCATCAGTAGACTCGAACTGTGAGACTATCTCGGTTATGACAGCAGAAATGCTCTTCTT 5278
Db 5085 GTCATCAGTAGACTCGAACTGTGAGACTATCTCGGTTATGACAGCAGAAATGCTCTTCTT 5144
Qy 5279 GGAGACAGTAATGAAGTCCCAACCAATAAGAAATCTTGTATTCAGGAACAAACTTCTT 5338
Db 5145 GGAGACAGTAATGAAGTCCCAACCAATAAGAAATCTTGTATTCAGGAACAAACTTCTT 5204
Qy 5339 GTTT--CGAACTTTTTCGCTGCTTGAACATAAATGTAGAGTGGATATGTCGGGTAGG 5396
Db 5205 GTTTGCGCAACTTTTTCGCTGCTTGAACATAAATGTAGAGTGGATATGTCGGGTAGG 5264
Qy 5397 AATGAGCGGGCAATGCTTACTTCTTGGACCTTCAAGAGGTATGTAGGTTTGTAGATA 5456
Db 5265 AATGAGCGGGCAATGCTTACTTCTTGGACCTTCAAGAGGTATGTAGGTTTGTAGATA 5324
Qy 5457 CTGATGCAACTTCAGTGACAAACGTTGCTATTTTCGTTTCAAAACCTTCCGAATCCAGAGAA 5516
Db 5325 CTGATGCAACTTCAGTGACAAACGTTGCTATTTTCGTTTCAAAACCTTCCGAATCCAGAGAA 5384
Qy 5517 ATCAAGTTGTTTGTCTACTATTTGATCAAGCCAGTCCGCTTGTGAACCTGACAAATAGTG 5576
Db 5385 ATCAAGTTGTTTGTCTACTATTTGATCAAGCCAGTCCGCTTGTGAACCTGACAAATAGTG 5444
Qy 5577 TGCTCGTCTTTGAGGTCATCTTTGATGAATAAATCTAGTCTTGTGATCTTAAATATCTT 5636
Db 5445 TGCTCGTCTTTGAGGTCATCTTTGATGAATAAATCTAGTCTTGTGATCTTAAATATCTT 5504
Qy 5637 GACGAGCCAGCGGATTAATCCCAATCTTAAATCTTCTTTAAACGTTTAAAGGCAAG 5696
Db 5505 GACGAGCCAGCGGATTAATCCCAATCTTAAACGTTTCTTAAAGGCAAG 5564
Qy 5697 TATGTCCTCTGATTAATCCCAATCTAGTCTGATCTGATCTTCACTCATCACTTGAAGG 5756
Db 5565 TATGTCCTCTGATTAATCCCAATCTAGTCTGATCTGATCTTCACTCATCACTTGAAGG 5624

QY 5757 GCATATCTGTTTATAGAAATTTGCGGAGATGCGATATCGAGAAAGGTACGCTGAT 5816
DB 5625 GCACATATCTGTTTATAGAAATTTGCGGAGATGCGATATCGAGAAAGGTACGCTGAT 5684
QY 5817 TTTAAACGTGAATTTATCTCAAGATCTCTGCTCG 5852
DB 5685 TTTAAACGTGAATTTATCTCAAGATCTCTGCTCGCGCG 5720

RESULT 14

US-10-038-722-70
; Sequence 70, Application US/10038722
; Publication No. US20030175919A1
; GENERAL INFORMATION:
; APPLICANT: LEY, Arthur C.
; APPLICANT: GUTERMAN, Sonia K.
; APPLICANT: MARKLAND, William
; APPLICANT: KENT, Rachel B.
; APPLICANT: ROBERTS, Bruce L.
; APPLICANT: LADNER, Robert C.
; TITLE OF INVENTION: ITI-D1 KUNITZ DOMAIN MUTANTS AS rHE INHIBITORS
; FILE REFERENCE: LEY-1B
; CURRENT APPLICATION NUMBER: US/10/038,722
; CURRENT FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 08/849,406
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: PCT/US95/16349
; PRIOR FILING DATE: 1995-12-15
; PRIOR APPLICATION NUMBER: US 08/358,160
; PRIOR FILING DATE: 1994-12-16
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70
; LENGTH: 8157
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Plasmid PHIL-D2 (Table 250)
US-10-038-722-70

Query Match 55.4%; Score 4565.2; DB 15; Length 8157;
Best Local Similarity 89.9%; Pred. No. 0;
Matches 5262; Conservative 0; Mismatches 13; Indels 579; Gaps 11;
QY 2 GATCTAACATCCAAAGAGAAAGGTTGAATGAACCTTTTGCCATCCGACATCCACAGG 61
DB 14 GATCTAACATCCAAAGAGAAAGGTTGAATGAACCTTTTGCCATCCGACATCCACAGG 73
QY 62 TCCATTCTCACATAAGTGCACAAACGCAACAGAGGGGATACACTAGCAGCAGACCGTT 121
DB 74 TCCATTCTCACATAAGTGCACAAACGCAACAGAGGGGATACACTAGCAGCAGACCGTT 133
QY 122 GCAACGCGAGGACCTCCACTCTCTCTCCCTCAACACCCACTTTTGCCATCGAAAACCA 181
DB 134 GCAACGCGAGGACCTCCACTCTCTCTCCCTCAACACCCACTTTTGCCATCGAAAACCA 193
QY 182 GCCCAGTTATTGGGCTTGATTTGGAGCTGGCTCATTCCTCAATTCCTTCTATTAGGCTACTAA 241
DB 194 GCCCAGTTATTGGGCTTGATTTGGAGCTGGCTCATTCCTCAATTCCTTCTATTAGGCTACTAA 253
QY 242 CACCATGACTTTATTAGCTGCTATTCCTGGCCCCCTTGGGAGGTTTATGTTTGTAT 301
DB 254 CACCATGACTTTATTAGCTGCTATTCCTGGCCCCCTTGGGAGG-TCATGTTTGTAT 312
QY 302 TTCCGAATCCAAAGCTCCGATTTACACCCGAAACATCACTCCAGATGAGGCTTTCTGA 361
DB 313 TTCCGAATCCAAAGCTCCGATTTACACCCGAAACATCACTCCAGATGAGGCTTTCTGA 372
QY 362 GTGTGGGTCAATAGTTTATGTTTCCCAATATGCCCCAAACTGACAGTTTAAACGCTG 421
DB 373 GTGTGGGTCAATAGTTTATGTTTCCCAATATGCCCCAAACTGACAGTTTAAACGCTG 431

QY 422 TCTTGGAACTTAATATGACAAAGCGTGATCTCATCAAGATGAACTAAGTTTGGTTCGT 481
DB 432 TCTTGGAACTTAATATGACAAAGCGTGATCTCATCAAGATGAACTAAGTTTGGTTCGT 491
QY 482 TGAATGCTAACCGCCAGTTTGGTCAAAAGAAACTTCCAAAAGTCGCGCATACCGTTTGTG 541
DB 492 TGAATGCTAACCGCCAGTTTGGTCAAAAGAAACTTCCAAAAGTCGCGCATACCGTTTGTG 551
QY 542 TTGTTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAATGCTTGAAGCGAGTCTC 601
DB 552 TTGTTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAATGCTTGAAGCGAGTCTC 611
QY 602 TCTATCGCTTCTGAAACCCCGGTGCACTGTGCGGAAACGCAAAATGGGGAAC-ACCCTGCT 660
DB 612 TCTATCGCTTCTGAAACCCCGGTGCGCACTGTGCGGAAACGCAAAATGGGGAAC-ACCCTGCT 671
QY 661 TTTTGGATGATTGCAATTTGT-CTCCACATTTGTATGCTTCAAGATTTCTGGTGGGAATAC 719
DB 672 TTTTGGATGATTGCAATTTGTCTCCACATTTGTATGCTTCAAGATTTCTGGTGGGAATAC 731
QY 720 TGCTGATAGCTTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTGACA-GCA 778
DB 732 TGCTGATAGCTTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTGACAGCA 791
QY 779 ATATATAACAGAGGAGGCTGCCCTGCTTAAACCTTTTATCATCATTTATAGC 838
DB 792 ATATATAACAGAGGAGGCTGCCCTGCTTAAACCTTTTATCATCATTTATAGC 851
QY 839 TTACTTTTCAATAATTTGCGACTGGTTTCCAATTTGACAAGCTTTTGAATTTTAAACGCTTTTAAAC 898
DB 852 TTACTTTTCAATAATTTGCGACTGGTTTCCAATTTGACAAGCTTTTGAATTTTAAACGCTTTTAAAC 911
QY 899 GACAACTTTGAGAAGATCAAAAACAACTAATTTTGAAGAGTCCAAAACGATGAGATTTTC 958
DB 912 GACAACTTTGAGAAGATCAAAAACAACTAATTTTGAAGAGTCCAAAACGATGAGATTTTC 953
QY 959 CTTCAATTTTACTGTCAGTTTATTTGCGAGCATCTCCGCAATTTAGCTGCTCCAGTCAACA 1018
DB 954 ----- 953
QY 1019 CTACAAGAGAGATGAACCGGCACAAATTCGGCTGAAGCTGTCATCGGTTACTCAGATT 1078
DB 954 ----- 953
QY 1079 TAGAAGGGATTTTCGATGTTGCTGTTTGCCATTTTCCACAGCACAATAACGGGTAT 1138
DB 954 ----- 953
QY 1139 TGTTTTAAATACTACTATTGTCAGCATTTGCTCTAAAGAAAGGGGTATCTCTCGAGA 1198
DB 954 ----- 953
QY 1199 AAAGAGGGCTGAAGCCCGCCAGGCCCGCCAGGCCCGCCAGGAGAAACAATGTATCA 1258
DB 954 ----- 953
QY 1259 TGGCCCCCGGGAGAGATTAATTTGGCTTCCCGGTTGTCACCGCCAGCAGTGCACGG 1318
DB 954 ----- 953
QY 1319 AGAGAGTTGCTGTTTGTATGACAGTGTCCGGGGATTCCTCGTGGTCTTCCACCCCATGG 1378
DB 954 ----- 953
QY 1379 CCATCGAGAACACTCAAGAAGAAGATGTCCTTCTTAAGTAGTGGCGTAGAATTCCTCTAG 1438
DB 954 ----- 953
QY 1439 GCGCGCGCGAATTAATTCGCTTTAGACATGACTGTTCTCAGTTCAAGTTGGGCACTTA 1498
DB 954 -----AGGAATTCGCTTTAGACATGACTGTTCTCAGTTCAAGTTGGGCACTTA 1501
QY 1499 CGAGAAGACCGGCTTGTCTAGATTCTTAATCAAGAGGATGTCAAGTGCCTTGTGAG 1558

Db 1002 CGAGAGACCGTCTTCTGCTAGATCTTAATCAGAGGATGTCAAGATGCCATTTGGCTCGAG 1061
Qy 1559 AGATGACAGGCTTCATTTTTCATGATCTTTTATTTTGAATCTATATAGTATAGATTTT 1618
Db 1062 AGATGACAGGCTTCATTTTTCATGATCTTTTATTTTGAATCTATATAGTATAGATTTT 1121
Qy 1619 TTGTCATTTTGTCTTCTCGTAGAGCTTGTCTGCTGATCAGCTATCTCGCAGCTGATG 1678
Db 1122 TTGTCATTTTGTCTTCTCGTAGAGCTTGTCTGCTGATCAGCTATCTCGCAGCTGATG 1181
Qy 1679 AATATCTTGTGTAGAGGTTTGGAAAAATCATTCGAGTTTGAATTTTCTTTGTTATTC 1738
Db 1182 AATATCTTGTGTAGAGGTTTGGAAAAATCATTCGAGTTTGAATTTTCTTTGTTATTC 1241
Qy 1739 CCACCTCTCTTCAGAGTACAGAGATTAAGTGAGAGTTCGTTGTGCAAGCTTATCGAT 1798
Db 1242 CCACCTCTCTTCAGAGTACAGAGATTAAGTGAGAGTTCGTTGTGCAAGCTTATCGAT 1301
Qy 1799 AAGCTTTTAAATCGGCTAGTTTATCACAGTTTAAATTCCTAAACGAGTCAGCACCGTGTATG 1858
Db 1302 AAGCTTTTAAATCGGCTAGTTTATCACAGTTTAAATTCCTAAACGAGTCAGCACCGTGTATG 1361
Qy 1859 AATATCTAACAAATGGCTTCATCGTCATCTCGGACCGCTCACCTCGGATGCTGTAGGATA 1918
Db 1362 AATATCTAACAAATGGCTTCATCGTCATCTCGGACCGCTCACCTCGGATGCTGTAGGATA 1421
Qy 1919 GGCTTGTGTTATGCGGTACTTGGCGGCTCTTGGCGGATPATGTCATTCGACAGCATC 1978
Db 1422 GGCTTGTGTTATGCGGTACTTGGCGGCTCTTGGCGGATPATGTCATTCGACAGCATC 1481
Qy 1979 GCAGTCACTATCGCTGCTGCTAGCGCTATATGCTGATGCTGATGCTGATGCTGATGCTG 2038
Db 1482 GCAGTCACTATCGCTGCTGCTAGCGCTATATGCTGATGCTGATGCTGATGCTGATGCTG 1541
Qy 2039 GTTCTCGGAGCACTGTCCGACCGCTTTGGCCGCGCCGACAGTCTGCTGCTGCTGCTGCT 2098
Db 1542 GTTCTCGGAGCACTGTCCGACCGCTTTGGCCGCGCCGACAGTCTGCTGCTGCTGCTGCT 1601
Qy 2099 GAGGACCTATCGACTACCGGATCTAGCGCTTACGCGACACACCGCTGCTGCTGCTGCTGCT 2158
Db 1602 GAGGACCTATCGACTACCGGATCTAGCGCTTACGCGACACACCGCTGCTGCTGCTGCTGCT 1661
Qy 2159 TAAATGTAAAGTTAAATCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2218
Db 1662 TAAATGTAAAGTTAAATCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1721
Qy 2219 AACAGCATTCGGGTAGAGCTATAGAGCTTCAACGACGCGAGATCCATCTGCTTGGCC 2278
Db 1722 AACAGCATTCGGGTAGAGCTATAGAGCTTCAACGACGCGAGATCCATCTGCTTGGCC 1781
Qy 2279 AATATGTTTCAGTCCCTCAGGAGTTACGTTCTGTAAGTGTAGTAACTTCTGGAAGTTGC 2338
Db 1782 AATATGTTTCAGTCCCTCAGGAGTTACGTTCTGTAAGTGTAGTAACTTCTGGAAGTTGC 1841
Qy 2339 AGTGTAACTCCGCTGTTATGAGGCGCATTCGTAAGTGTGCAAGTGTGCTGCTGCTGCT 2398
Db 1842 AGTGTAACTCCGCTGTTATGAGGCGCATTCGTAAGTGTGCAAGTGTGCTGCTGCTGCTGCT 1901
Qy 2399 GAGGAGTAACTTCCAACTCTCTGGAGAGTAGGACCAACAAACAGATCCAGCGGTG 2458
Db 1902 GAGGAGTAACTTCCAACTCTCTGGAGAGTAGGACCAACAAACAGATCCAGCGGTG 1961
Qy 2459 TTGTACTCTGATCAACATAGAGAGAGCATTCGATTTGCAAGGATCAAGTGTTCAGGAGC 2518
Db 1962 TTGTACTCTGATCAACATAGAGAGAGCATTCGATTTGCAAGGATCAAGTGTTCAGGAGC 2021
Qy 2519 GTACTGATTTGGAATTTCCAAAGCTGCTCGTAGGTTGCAACCGATAGGTTGTAGAGTG 2578
Db 2022 GTACTGATTTGGAATTTCCAAAGCTGCTCGTAGGTTGCAACCGATAGGTTGTAGAGTG 2081
Qy 2579 TGCATATACACTTGGGTACAAATTTCAACCTTGGCACTGCAAGCTTGGTGTGCAAGC 2638

Db 2082 TGCATATACACTTGGGTACAAATTTCAACCTTGGCACTGCAAGCTTGGTGTGCAAGC 2141
Qy 2639 ATCTTCAATTTCTGGCAAGCTCTCTCTGTCTGTATATATCGACGCAACAGAAATCATCTGGGA 2698
Db 2142 ATCTTCAATTTCTGGCAAGCTCTCTCTGTCTGTATATATCGACGCAACAGAAATCATCTGGGA 2201
Qy 2699 ATCAATACCATCTTTCAGCTTGTAGAGCAGAGGCTCTGAGGCAACGAAATCTGATCAGCGTA 2758
Db 2202 ATCAATACCATCTTTCAGCTTGTAGAGCAGAGGCTCTGAGGCAACGAAATCTGATCAGCGTA 2860
Qy 2759 TTTATCAGCAATTAATCTAGAACTTTCAGAGGCCCCAGGCAATGTCAATPATCTACACGCGC 2818
Db 2261 TTTATCAGCAATTAATCTAGAACTTTCAGAGGCCCCAGGCAATGTCAATPATCTACACGCGC 2320
Qy 2819 TGATGTGTCTTTCGAAACCATCATCTTTCGAGCAGCTTACGAACTGGTTTCTCGGACCAAA 2878
Db 2321 TGATGTGTCTTTCGAAACCATCATCTTTCGAGCAGCTTACGAACTGGTTTCTCGGACCAAA 2380
Qy 2879 TATTTGTCTACATTTAGGAACAGTTTCTTCCGTAAGCCATAGCACTACTGCTCGGC 2938
Db 2381 TATTTGTCTACATTTAGGAACAGTTTCTTCCGTAAGCCATAGCACTACTGCTCGGC 2440
Qy 2939 GCCTCTCTGACGACGATACATCTTTCAGCACCACCTTTCGGGCAACGTAAGTCTCTGG 2998
Db 2441 GCCTCTCTGACGACGATACATCTTTCAGCACCACCTTTCGGGCAACGTAAGTCTCTGG 2500
Qy 2999 GGTAAAGGTTACCATCTTCTTCTAGTGGAGATGCAAAAAACAAATTTCTTTCGCAACGCAAC 3058
Db 2501 GGTAAAGGTTACCATCTTCTTCTAGTGGAGATGCAAAAAACAAATTTCTTTCGCAACGCAAC 2560
Qy 3059 TTTGCAAGGAAACACCCAGCATCTGAGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3118
Db 2561 TTTGCAAGGAAACACCCAGCATCTGAGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2820
Qy 3119 ATAGAGGCAACTTTCTCAATAGTCTTTCGCAAAACGAGAGCAGACTACACAGGCGCAAGT 3178
Db 2621 ATAGAGGCAACTTTCTCAATAGTCTTTCGCAAAACGAGAGCAGACTACACAGGCGCAAGT 2680
Qy 3179 CTCACATTCGACGCTCTCGTCTGAGCTTCTATGGAATTCCTGAGAGTATCTATAGA 3238
Db 2681 CTCACATTCGACGCTCTCGTCTGAGCTTCTATGGAATTCCTGAGAGTATCTATAGA 2740
Qy 3239 GAGATCAATGCTCTTTCATAGCTTATCTGCAATTCGCAATTCCTCTGGAAGAGGAGC 3298
Db 2741 GAGATCAATGCTCTTTCATAGCTTATCTGCAATTCGCAATTCCTCTGGAAGAGGAGC 2800
Qy 3299 TTTTCAACAGAGTGTCTTCAAGGAGCTCCATCAAACTTGGCAGTGTAGTCTTAAAGGGC 3358
Db 2801 TTTTCAACAGAGTGTCTTCAAGGAGCTCCATCAAACTTGGCAGTGTAGTCTTAAAGGGC 2860
Qy 3359 TTTTGTCACTTTTGAAGCAATTCGCAATTCGCAATTCGCAATTCGCAATTCGCAATTCGCAAT 3418
Db 2861 TTTTGTCACTTTTGAAGCAATTCGCAATTCGCAATTCGCAATTCGCAATTCGCAATTCGCAAT 2920
Qy 3419 GGTTTTCTGATAGGAGCAGGAGGAGCTTCAAACTTTCGAGGAGGAGGAGGAGGAGGAGGAGGAG 3478
Db 2921 GGTTTTCTGATAGGAGCAGGAGGAGCTTCAAACTTTCGAGGAGGAGGAGGAGGAGGAGGAGGAG 2980
Qy 3479 GTCAATTTTTCGACAAATTCATACAGCTTTCAGAGGAGGAGCTTCTTCTAGGTTTGGATCTTTC 3538
Db 2981 GTCAATTTTTCGACAAATTCATACAGCTTTCAGAGGAGGAGCTTCTTCTAGGTTTGGATCTTTC 3040
Qy 3539 TTTTAGGTTTCTTCTGATGATCTTCTGCTTGGGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3598
Db 3041 TTTTAGGTTTCTTCTGATGATCTTCTGCTTGGGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3100
Qy 3599 GGACTTCATATCCAGGTTTCTTCTCACCTCTGTCACAGTCCACAGTCCACAGTCCACAGTCCACAGT 3658
Db 3101 GGACTTCATATCCAGGTTTCTTCTCACCTCTGTCACAGTCCACAGTCCACAGTCCACAGTCCACAGT 3160
Qy 3659 AACTAATTCGCAAAATTAATTAAGTACAGCAGATTCGAGGCTATATCTTCTTCTTCTTCTTCTTCT 3718
Db 3161 AACTAATTCGCAAAATTAATTAAGTACAGCAGATTCGAGGCTATATCTTCTTCTTCTTCTTCTTCT 3220

3719 TTCTGCAAGTTTCATCAGCTTCCTCCCTAAATTTTAGCGTTCAACAAAACCTTCGTGCTCAAA 3778
3221 TTCTGCAAGTTTCATCAGCTTCCTCCCTAAATTTTAGCGTTCAACAAAACCTTCGTGCTCAAA 3280
3779 TAACCGTTTGGTATAGAACCTTCGGAGCATTTGCTTTACGATCCGACAAAGTGGCTTC 3838
3281 TAACCGTTTGGTATAGAACCTTCGGAGCATTTGCTTTACGATCCGACAAAGT-GCTTC 3339
3839 CATGGCTTAAGACCTTTGATTGGCCAAAACAGGAAGTGCCTTCCAAAGTGACAGAAACC 3898
3340 CATGGCTTAAGACCTTTGATTGGCCAAAACAGGAAGTGCCTTCCAAAGTGACAGAAACC 3399
3899 AACACCTGTTTGTCAACACAAAATTTCAAGCAGTCTCCATCACAATCCAATTCGATACC 3958
3400 AACACCTGTTTGTCAACACAAAATTTCAAGCAGTCTCCATCACAATCCAATTCGATACC 3459
3959 CAGCAACTTTTCAAGTGTCTCCAGATGTAGACCTTTTATACCAACAAACCGTGACGACGAGA 4018
3460 CAGCAACTTTTCAAGTGTCTCCAGATGTAGACCTTTTATACCAACAAACCGTGACGACGAGA 3519
4019 TTGGTAGATCCAGTTTGTCTTATAGCTCCGGAATPAGACTTTTGGACGAGTACAC 4078
3520 TTGGTAGATCCAGTTTGTCTTATAGCTCCGGAATPAGACTTTTGGACGAGTACAC 3579
4079 CAGGCCAACGAGTAATTAGAGAGTCAAGCACCAGTAAGTGTGATAGACCATCGGGCG 4138
3580 CAGGCCAACGAGTAATTAGAGAGTCAAGCACCAGTAAGTGTGATAGACCATCGGGCG 3639
4139 GTCAAGTCAAGAGCGCCAAACAAAATTTCACTGACAGGAACTTTTGAATCTTCA 4198
3640 GTCAAGTCAAGAGCGCCAAACAAAATTTCACTGACAGGAACTTTTGAATCTTCA 3699
4199 AAGTTGCTATTCAAGTCAATTTGGGAGCATCAATATAGGGATTTATACAGAAAGCAAC 4258
3700 AAGTTGCTATTCAAGTCAATTTGGGAGCATCAATATAGGGATTTATACAGAAAGCAAC 3759
4259 AGTGAAGTCAATCTACCAACTTTTGGGAGCATCAATATAGGGATTTATACAGAAAGCAAC 4318
3760 AGTGAAGTCAATCTACCAACTTTTGGGAGCATCAATATAGGGATTTATACAGAAAGCAAC 3819
4319 GCATTAAGTCAATCTTCAATTTCCGAGTGGAGAAAGCAAGCAAGCACTACTAGC 4378
3820 GCATTAAGTCAATCTTCAATTTCCGAGTGGAGAAAGCAAGCAAGCACTACTAGC 3879
4379 ATTAGGGCAGAGTGAATTTATCAACAGGCTCTATAGATAACCCCTAGCGCTGG 4438
3880 ATTAGGGCAGAGTGAATTTATCAACAGGCTCTATAGATAACCCCTAGCGCTGG 3939
4439 GATCATCTTTGGCAAACTCTTTCTGCCAAATCTAGGTCCTCAAAATCACTTCAATGATACC 4498
3940 GATCATCTTTGGCAAACTCTTTCTGCCAAATCTAGGTCCTCAAAATCACTTCAATGATACC 3999
4499 ATTATGTACAACTTGAGCAAGTGTGCGATCAGCTCCTCAAAATGCTGCTGTAACGGA 4558
4000 ATTTATTTT-----ACGGA 4009
4559 TGACTCAACTTGCAATTAATTTGAAGTCACTGATGAGTGAAGTCACTTGAAGTGGTGG 4618
4010 TGACTCAACTTGCAATTAATTTGAAGTCACTGATGAGTGAAGTCACTTGAAGTGGTGG 4069
4619 CAGCTGGTCAGGATAGGGAACACCGCTTTTCTTCAACAACTCAAGGAATTTATCAAA 4678
4070 CAGCTGGTCAGGATAGGGAACACCGCTTTTCTTCAACAACTCAAGGAATTTATCAAA 4129
4679 CTCTCAACACTTGGTATGAGGTAGCAAGGGAATGTCATCTTGAAGTGGGACAGTG 4738
4130 CTCTCAACACTTGGTATGAGGTAGCAAGGGAATGTCATCTTGAAGTGGGACAGTG 4189
4739 AGTGTAGTCTTGAAGAAATTTCTGAAGCGGTATTTTATTTATCAGTCACTGATCAGGA 4798
4190 AGTGTAGTCTTGAAGAAATTTCTGAAGCGGTATTTTATTTATCAGTCACTGATCAGGA 4249

4799 GATCCTCTAGCCGAGCGCATCGTGGCCGACCTCGTAGGTGGCATCACCGGCGCCACAGG 4858
4250 GATCCTCTAGCCGAGCGCATCGTGGC-----CGGCATCACCGGCGCCACAGG 4297
4859 TGGGTTTGGTGGCCCTTATATCCCGGACATCACCGATGGGGAAGATCGGGCTGCGCACTT 4918
4298 TGGGTTTGGTGGGCGCTTATATCCCGGACATCACCGATGGGGAAGATCGGGCTGCGCACTT 4357
4919 CGGGCTCATGAGCGCTTTTGGCGCTGGGTATGGTGGCAGGCGCCCTGGCGGGGACT 4978
4358 CGGGCTCATGAGCGCTTTTGGCGCTGGGTATGGTGGCAGGCGCCCTGGCGGGGACT 4417
4979 GTTGGGCGCCATCTCTTTGATGACCAATCTTTCGGCGCGGCGGTGCTCAACGGCCTCAA 5038
4418 GTTGGGCGCCATCTCTTTGATGACCAATCTTTCGGCGCGGCGGTGCTCAACGGCCTCAA 4477
5039 CCTACTACTGGGCTGCTTCTTAATGAGGAGTCCGATAAGGAGAGCGTCAAGTATCTAT 5098
4478 CCTACTACTGGGCTGCTTCTTAATGAGGAGTCCGATAAGGAGAGCGTCAAGTATCTAT 4537
5099 GATTGGAAGTATGGGAATGGTATACCCGCAATCTTTCAGTGTCTTGAAGTCTCTATCAG 5158
4538 GATTGGAAGTATGGGAATGGTATACCCGCAATCTTTCAGTGTCTTGAAGTCTCTATCAG 4597
5159 ATTATGCCCAACTAAAGCAACCGGAGGAGAGATTTTCATGGTAAATTTCTCTGACTTTTG 5218
4598 ATTATGCCCAACTAAAGCAACCGGAGGAGAGATTTTCATGGTAAATTTCTCTGACTTTTG 4657
5219 GTCATCAGTAGACTCGAATCTGTGAGACTATCTCCGTTATCACAGCAGAAATGTCCTTCTT 5278
4658 GTCATCAGTAGACTCGAATCTGTGAGACTATCTCCGTTATCACAGCAGAAATGTCCTTCTT 4717
5279 GAGAGCAGTAAATGAAGTCCCAACAAATAAGAAATCTTGTATACAGAAACAACTTCTT 5338
4718 GAGAGCAGTAAATGAAGTCCCAACAAATAAGAAATCTTGTATACAGAAACAACTTCTT 4777
5339 GTTTCGAACTTTTTCGGTGCCTTGAACATAAAATGTAGAGTGGATATGTCGGGTAGGAA 5398
4778 GTTTCGAACTTTTTCGGTGCCTTGAACATAAAATGTAGAGTGGATATGTCGGGTAGGAA 4837
5399 TGGAGGGGCAATGCTTACCTTCTGACCTTCAAGAGTATGATAGGTTTGTAGATCT 5458
4838 TGGAGGGGCAATGCTTACCTTCTGACCTTCAAGAGTATGATAGGTTTGTAGATCT 4897
5459 GATGCCAACTTCAGTGACAGCTTATTTTCGTTCAAAACCATTCGAAATCCAGAGAAAT 5518
4898 GATGCCAACTTCAGTGACAGCTTATTTTCGTTCAAAACCATTCGAAATCCAGAGAAAT 4957
5519 CAAAGTTGTTTGTCTACTATTGATCCAAAGCCAGTGGGCTTTCGAAACTGACAAATAGTGTG 5578
4958 CAAAGTTGTTTGTCTACTATTGATCCAAAGCCAGTGGGCTTTCGAAACTGACAAATAGTGTG 5017
5579 CTGCTGTTTGGAGTCACTTTTGTATGAATAAATCTAGTCTTGTATCTAAATATCTTGA 5638
5018 CTGCTGTTTGGAGTCACTTTTGTATGAATAAATCTAGTCTTGTATCTAAATATCTTGA 5077
5639 CGAGCCAAAGCGATAAATACCCAAATCTTAAACCTTTTAAACGTTTAAAGGACAAAGTA 5698
5078 CGAGCCAAAGCGATAAATACCCAAATCTTAAACCTTTTAAACGTTTAAAGGACAAAGTA 5137
5699 TGTCTGCTGATAAACCACCAATCAGCTCGTGTAGTCTGATCTCTATCAACTTGTAGGGGC 5758
5138 TGTCTGCTGATAAACCACCAATCAGCTCGTGTAGTCTGATCTCTATCAACTTGTAGGGGC 5197
5759 ACTATCTGTTTGTAGAGAAATTTGGGAGATCGGATATCGAGAAAGGTACGCTGATTT 5818
5198 ACTATCTGTTTGTAGAGAAATTTGGGAGATCGGATATCGAGAAAGGTACGCTGATTT 5818
5819 TAAACGTGAAATTTATCTCAAGATCTCTGCCTCG 5852
5258 TAAACGTGAAATTTATCTCAAGATCGGCGCGCG 5291

RESULT 15

US-10-179-046-19/c
; Sequence 19, Application US/10179046
; Publication No.: US20030031354A1
; GENERAL INFORMATION:
; APPLICANT: Crawford, Kenneth
; Zaror, Isahel
; Innis, Michael
; TITLE OF INVENTION: Pichia Secretory Leader for Protein
; Expression
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: United States
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10179,046
; FILING DATE: 25-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/029,267
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Chung, Ling-Feng
; REGISTRATION NUMBER: 36,482
; REFERENCE/DOCKET NUMBER: 1165.100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2704
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; TYPE: nucleic acid
; LENGTH: 2659 base pairs
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-10-179-046-19

Query Match 32.3%; Score 2659; DB 15; Length 2659;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2145 GGATCTATCGAATCTAAATGTAAGTTAAATCTCTAAATAATTAATAAGTCCAGTTTC 2204
DB 2659 GGATCTATCGAATCTAAATGTAAGTTAAATCTCTAAATAATTAATAAGTCCAGTTTC 2600
QY 2205 TCCATACGACCTTAACAGCATGGGTGAGCATCTAGACCTTCAACAGAGCCAGATCC 2264
DB 2599 TCCATACGACCTTAACAGCATGGGTGAGCATCTAGACCTTCAACAGAGCCAGATCC 2540
QY 2265 ATCACTGCTTGGCCCAATATGTTTCAGTCCCTCAGGAGTTACGTCTTGTGAAGTCAATGAAC 2324
DB 2539 ATCACTGCTTGGCCCAATATGTTTCAGTCCCTCAGGAGTTACGTCTTGTGAAGTCAATGAAC 2480
QY 2325 TTCTGGAAGTTGCAAGTGTACTCGGTGATTCACGGGCATATCGTACGTTGGCAAA 2384
DB 2479 TTCTGGAAGTTGCAAGTGTACTCGGTGATTCACGGGCATATCGTACGTTGGCAAA 2420
QY 2385 GTGTGTTGGTACCGGAGGAGTAATCTCCAACTCTCTCTGAGAGTAGGCAACCAACAAAC 2444
DB 2419 GTGTGTTGGTACCGGAGGAGTAATCTCCAACTCTCTCTGAGAGTAGGCAACCAACAAAC 2360
QY 2445 ACAGATCCAGCGTGTGATCTGATCAACATAGAGAGCATCTCGATTGCGAGATC 2504

DB 2359 ACAGATCCAGCGTGTGATCTGATCAACATAAGAGAGCAATTCGATTCGAGGATC 2300
QY 2505 AAGTGTTCAGGAGCGTACTGATTTGACATTTCCAAAGCCCTGCTGTAGTTGCAACCGAT 2564
DB 2299 AAGTGTTCAGGAGCGTACTGATTTGACATTTCCAAAGCCCTGCTGTAGTTGCAACCGAT 2240
QY 2565 AGGTGTGTAGAGTGTCAATACATCTGCGTACAAATTTCAACCCCTTGGCAACTGCACAGCT 2624
DB 2239 AGGTGTGTAGAGTGTCAATACATCTGCGTACAAATTTCAACCCCTTGGCAACTGCACAGCT 2180
QY 2625 TGGTGTGTGACAGCATCTTCAATTTGCGAAGCTCTTGTCTGTCTATATCGACAGCCAAAC 2684
DB 2179 TGGTGTGTGACAGCATCTTCAATTTGCGAAGCTCTTGTCTGTCTATATCGACAGCCAAAC 2120
QY 2685 AGAATCACCTGGGAATCAATACCATGTTTCACTTTGAGACAGAAAGTCTGAGGCAACGAAA 2744
DB 2119 AGAATCACCTGGGAATCAATACCATGTTTCACTTTGAGACAGAAAGTCTGAGGCAACGAAA 2060
QY 2745 TCTGATCAGCGTATTTATCAGCAATTAATAGAACTTCAGAGGCCCGAGGAGCATGTCA 2804
DB 2059 TCTGATCAGCGTATTTATCAGCAATTAATAGAACTTCAGAGGCCCGAGGAGCATGTCA 2000
QY 2805 ATACTACACAGGCGTGTGATGTTTGAACCATCATCTTGGCAGCAGTAACGAACTGG 2864
DB 1999 ATACTACACAGGCGTGTGATGTTTGAACCATCATCTTGGCAGCAGTAACGAACTGG 1940
QY 2865 TTTCTGTGACCAATATTTTGTCACTTAGGAACAGTTTCTGTTCGTAAGCCATAGCA 2924
DB 1939 TTTCTGTGACCAATATTTTGTCACTTAGGAACAGTTTCTGTTCGTAAGCCATAGCA 1880
QY 2925 GCTACTGCTGGGCGCTCTCTAGCAGCATACATTTAGCACCACCTTTGTGGGCAACG 2984
DB 1879 GCTACTGCTGGGCGCTCTCTAGCAGCATACATTTAGCACCACCTTTGTGGGCAACG 1820
QY 2985 TAGATGACTTCTGGGGTAAGGGTACCATTCTTTAGTGGAGATGCAAAAACAATTTCT 3044
DB 1819 TAGATGACTTCTGGGGTAAGGGTACCATTCTTTAGTGGAGATGCAAAAACAATTTCT 1760
QY 3045 TTGCAACAGCAACTTTTGGCAGGAACACCCAGCATCAGGAAAGTGGAAAGCAGAAATTCG 3104
DB 1759 TTGCAACAGCAACTTTTGGCAGGAACACCCAGCATCAGGAAAGTGGAAAGCAGAAATTCG 1700
QY 3105 GTTCCACAGGAATATAGAGCCCAAATTTCTCAATAGGTCTTGCAAAACAGAGAGCAGCT 3164
DB 1699 GTTCCACAGGAATATAGAGCCCAAATTTCTCAATAGGTCTTGCAAAACAGAGAGCAGCT 1640
QY 3165 ACACAGGCGCAAGTCTCAACTTGCAAGCTCTCGTTAGTTGAGCTTCATGGAATTTCCCTG 3224
DB 1639 ACACAGGCGCAAGTCTCAACTTGCAAGCTCTCGTTAGTTGAGCTTCATGGAATTTCCCTG 1580
QY 3225 AGTTTATCTATAGAGAGATCAATGGCTCTCTTAACTGTTATCTGGCAATTCGATAAGTTCC 3284
DB 1579 AGTTTATCTATAGAGAGATCAATGGCTCTCTTAACTGTTATCTGGCAATTCGATAAGTTCC 1520
QY 3285 TCTGGAAAGAGAGCTCTTAACACAGAGTGTCTTAAAGCGACTCCATCAAACTTTGGCAGTT 3344
DB 1519 TCTGGAAAGAGAGCTCTTAACACAGAGTGTCTTAAAGCGACTCCATCAAACTTTGGCAGTT 1460
QY 3345 AGTTCTAAAAGGGCTTGTGTACCATTTTGAACGAACTTGTGCAAAATTTGTTGACTAAT 3404
DB 1459 AGTTCTAAAAGGGCTTGTGTACCATTTTGAACGAACTTGTGCAAAATTTGTTGACTAAT 1400
QY 3405 TCCATAATCTGTTCGTTTCTGGATAGGACGAGAGGCACTTCAATTTCTGTGTAG 3464
DB 1399 TCCATAATCTGTTCGTTTCTGGATAGGACGAGAGGCACTTCAATTTCTGTGTAG 1340
QY 3465 GAGGCTTTAGAAAAGCTCAATTTTGCACAAATTCACACCTTCAGAGGAGCTTTCTTTA 3524
DB 1339 GAGGCTTTAGAAAAGCTCAATTTTGCACAAATTCACACCTTCAGAGGAGCTTTCTTTA 1280
QY 3525 GGTTTGGAATCTCTTTAGTGTGTTCCCTGTGCTATCTCGCTTGGCATCTCTCTTTCCCTT 3584
DB 1279 GGTTTGGAATCTCTTTAGTGTGTTCCCTGTGCTATCTCGCTTGGCATCTCTCTTTCCCTT 1220

```
QY 3585 CTAGTGACCTTTAGGACCTTCATATCCAGGTTTCTCTCAACCTCGTCCAACTGCACACG 3644
Db |||||
QY 1219 CTAGTGACCTTTAGGACCTTCATATCCAGGTTTCTCTCAACCTCGTCCAACTGCACACG 1160
Db |||||
QY 3645 TACTTGGCACAATCTAATAATGCAAAATAAATAAGTCAGACATTTCCAGGCTATATCT 3704
Db |||||
QY 1159 TACTTGGCACAATCTAATAATGCAAAATAAATAAGTCAGACATTTCCAGGCTATATCT 1100
Db |||||
QY 3705 TCCTTGGATTTAGCTTCTGCAAGTTTCATCAGCTTCTCTCCCTAAATTTAGCGTTCAACAAA 3764
Db |||||
QY 1099 TCCTTGGATTTAGCTTCTGCAAGTTTCATCAGCTTCTCTCCCTAAATTTAGCGTTCAACAAA 1040
Db |||||
QY 3765 ACTTCTCGTCAAAATAACCGTTTGGTATAGAAACCTTTCTGGAGCATTTGCTTTACGATCC 3824
Db |||||
QY 1039 ACTTCTCGTCAAAATAACCGTTTGGTATAGAAACCTTTCTGGAGCATTTGCTTTACGATCC 980
Db |||||
QY 3825 CACAAGTGGCTTCCATGGCTCTAAGACCCCTTTGATTTGGCCAAAACAGGAAGTGGCTTCC 3884
Db |||||
QY 979 CACAAGTGGCTTCCATGGCTCTAAGACCCCTTTGATTTGGCCAAAACAGGAAGTGGCTTCC 920
Db |||||
QY 3885 AAGTGACAGAAACCAACACCTGTTTGTTCACACCAAAATTTCAAGCAGTCTCCATCACA 3944
Db |||||
QY 919 AAGTGACAGAAACCAACACCTGTTTGTTCACACCAAAATTTCAAGCAGTCTCCATCACA 860
Db |||||
QY 3945 TCCAATTCGATACCCAGCAACTTTTGAGTTGCTCCAGATGTAGCACCTTTATACCACAAA 4004
Db |||||
QY 859 TCCAATTCGATACCCAGCAACTTTTGAGTTGCTCCAGATGTAGCACCTTTATACCACAAA 800
Db |||||
QY 4005 CCGTGACGACGAGATTTGGTAGTCTCCAGTTTGTGCTTATAGCTTCCGGATAGACTTT 4064
Db |||||
QY 799 CCGTGACGACGAGATTTGGTAGTCTCCAGTTTGTGCTTATAGCTTCCGGATAGACTTT 740
Db |||||
QY 4065 TTGGACGAGTACACCAAGGCCCAACGAGTAATTTAGAGAGTCAGCCACCAAAAGTAGTGAAT 4124
Db |||||
QY 739 TTGGACGAGTACACCAAGGCCCAACGAGTAATTTAGAGAGTCAGCCACCAAAAGTAGTGAAT 680
Db |||||
QY 4125 AGACCAATCGGGGGTCACTAGTCAAGACGCGCAACAAATTTCTGACAGGGAACCTTT 4184
Db |||||
QY 679 AGACCAATCGGGGGTCACTAGTCAAGACGCGCAACAAATTTCTGACAGGGAACCTTT 620
Db |||||
QY 4185 TTGACATCTTCAGAAAGTTTCGATTCAGTAGTCAATTTGCCGAGCATCAATATGGGGATT 4244
Db |||||
QY 619 TTGACATCTTCAGAAAGTTTCGATTCAGTAGTCAATTTGCCGAGCATCAATATGGGGATT 560
Db |||||
QY 4245 ATACCAAGAACCAAGTGAAGTCAATCTACAACTTTTGGGTTCTCAGAAAAAGCATAA 4304
Db |||||
QY 559 ATACCAAGAACCAAGTGAAGTCAATCTACAACTTTTGGGTTCTCAGAAAAAGCATAA 500
Db |||||
QY 4305 ACAGTTCTACTACCGCCATTAGTGAAACTTTTCAATCGCCAGTGGAGAGAAAAAGGC 4364
Db |||||
QY 499 ACAGTTCTACTACCGCCATTAGTGAAACTTTTCAATCGCCAGTGGAGAGAAAAAGGC 440
Db |||||
QY 4365 ACAGGATACTAGCATTAGCGGGCAAGGATGCAACTTTTATCAACGAGGTCCTATAGATA 4424
Db |||||
QY 439 ACAGGATACTAGCATTAGCGGGCAAGGATGCAACTTTTATCAACGAGGTCCTATAGATA 380
Db |||||
QY 4425 ACCCTAGCGCTGGGATCATCTTTGGACAATCTTTTCCCAAAATCTAGTCCAAAATC 4484
Db |||||
QY 379 ACCCTAGCGCTGGGATCATCTTTGGACAATCTTTTCCCAAAATCTAGTCCAAAATC 320
Db |||||
QY 4485 ACTTCATTGATACCAATTTAGTACAACTTTGAGCAAGTTGTGATCAGCTCCTCAAAATGG 4544
Db |||||
QY 319 ACTTCATTGATACCAATTTAGTACAACTTTGAGCAAGTTGTGATCAGCTCCTCAAAATGG 260
Db |||||
QY 4545 TCCTCTGTAACGGATGACTCAACTTGCACATTTAAGCTCAGTCAAGTGAAC 4604
Db |||||
QY 259 TCCTCTGTAACGGATGACTCAACTTGCACATTTAAGCTCAGTCAAGTGAAC 200
Db |||||
QY 4605 TTGATCAGGTTGTGAGCTGTGACGACATAGGGAACACAGGCTTTTCTACCAACTC 4664
Db |||||
QY 199 TTGATCAGGTTGTGAGCTGTGACGACATAGGGAACACAGGCTTTTCTACCAACTC 140
Db |||||
```

```
QY 4665 AAGGAATTATCAAACTCTGCAACACTTGCCTATGTCAGGTAGCAGGGAATGTCTACTTT 4724
Db |||||
QY 139 AAGGAATTATCAAACTCTGCAACACTTGCCTATGTCAGGTAGCAGGGAATGTCTACTTT 80
Db |||||
QY 4725 GAAGTCGACAGTCAGTGTAGTCTTTGAGAAATTTCTGAAGCCGTATTTTATTCAGTGA 4784
Db |||||
QY 79 GAAGTCGACAGTCAGTGTAGTCTTTGAGAAATTTCTGAAGCCGTATTTTATTCAGTGA 20
Db |||||
QY 4785 GTCAGTCAATCAGGAGATCC 4803
Db |||||
QY 19 GTCAGTCAATCAGGAGATCC 1
```

Search completed: September 24, 2004, 19:35:30
Job time : 2991.74 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2004, 18:06:49 ; Search time 352.227 Seconds
(without alignments)
12984.105 Million cell updates/sec

Title: US-10-030-390-3

Perfect score: 841

Sequence: 1 agatctaacatccaagacg.....aaataacagttattatcg 8241

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5142.8	62.4	8584	1	US-08-358-160-66
2	5130.8	62.3	8590	1	US-08-358-160-70
3	4565.2	55.4	8157	1	US-08-358-160-65
4	2659	32.3	2659	3	US-09-029-367-19
5	2310.8	28.0	34303	2	US-08-735-609-4
6	2310.8	28.0	34303	2	US-08-735-609-4
7	2310.8	28.0	34303	3	US-09-315-372-4
8	2310.8	28.0	34303	3	US-09-244-752-4
9	2310.8	28.0	34303	3	US-09-245-497-4
10	2310.8	28.0	34303	4	US-09-562-919-4
11	2302	27.9	6727	2	US-08-125-462-2
12	2302	27.9	6727	2	US-08-891-848-2
13	2302	27.9	6727	2	US-08-125-462-5
14	2302	27.9	6799	2	US-08-851-848-5
15	2302	27.9	8491	2	US-08-757-439-1
16	2301.4	27.9	8491	2	US-08-470-881-1
17	2299.6	27.9	5438	3	US-08-809-326A-25
18	2299.6	27.9	5438	4	US-09-914A-25
19	2299.6	27.9	5438	4	US-09-689-913A-25
20	2299.6	27.9	5438	4	US-09-689-916A-25
21	2299.6	27.9	5658	3	US-08-809-326A-10
22	2299.6	27.9	5658	4	US-09-689-914A-10
23	2299.6	27.9	5658	4	US-09-689-913A-10
24	2299.6	27.9	5658	4	US-09-689-916A-10
25	2299.6	27.9	5825	4	US-08-809-513A-7
26	2298.8	27.9	8068	4	US-09-301-593-27
27	2298.8	27.9	8068	4	US-09-301-593-35

Sequence 1, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 40, Appli
Sequence 2, Appli
Sequence 3, Appli
Sequence 14, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 3, Appli
Sequence 78, Appli
Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-08-358-160-66

; Sequence 66, Application US/08358160

; Patent No. 5663143

; GENERAL INFORMATION:

; APPLICANT: LEY, Arthur C.

; APPLICANT: LADNER, Robert C.

; APPLICANT: GUTERMAN, Sonia K.

; APPLICANT: ROBERTS, Bruce L.

; APPLICANT: MARKLAND, William

; APPLICANT: KENT, Rachel B.

; TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ

; NUMBER OF SEQUENCES: 234

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BROWDY AND NEIMARK

; STREET: 419 Seventh Street, N.W. Suite 300

; CITY: Washington

; STATE: District of Columbia

; COUNTRY: USA

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/358,160

; FILING DATE: 15-DEC-1994

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/133,031

; FILING DATE: 13-OCT-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/009,319

; FILING DATE: 26-JAN-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/664,989

; FILING DATE: 01-MAR-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/487,063

; FILING DATE: 02-MAR-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/240,160

; FILING DATE: 02-SEP-1988

; ATTORNEY/AGENT INFORMATION:

; NAME: Cooper, Iver P.

; REGISTRATION NUMBER: 28,005

; REFERENCE/DOCKET NUMBER: LEY=1

; TELECOMMUNICATION INFORMATION:

Mon Sep 27 14:45:39 2004

TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633

INFORMATION FOR SEQ ID NO: 66:

SEQUENCE CHARACTERISTICS:

LENGTH: 8584 bases

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: circular

MOLECULE TYPE: other nucleic acid

DESCRIPTION: DNA plasmid

US-08-358-160-66

Query Match 62.4%; Score 5142.8; DB 1; Length 8584;

Best Local Similarity 94.6%; Pred. No. 0;

Matches 5535; Conservative 0; Mismatches 167; Indels 152; Gaps 14;

Qy	2	GATCTAACATCCAAAGACGAAAGTTCGAATGAAACCTTTTGGCCATCCGACATCCACAGG	61
Db	14	GATCTAACATCCAAAGACGAAAGTTCGAATGAAACCTTTTGGCCATCCGACATCCACAGG	73
Qy	62	TCCATTCTCACATTAAGTGCMAACGCAACGAGGAGGATACACTAGCAGCAGACCGTT	121
Db	74	TCCATTCTCACATTAAGTGCMAACGCAACGAGGAGGATACACTAGCAGCAGACCGTT	133
Qy	122	GCAAAACGAGGACCTCCACTCTCTCTCTCAACCCCACTTTTGGCCATCCGAAACCA	181
Db	134	GCAAAACGAGGACCTCCACTCTCTCTCTCAACCCCACTTTTGGCCATCCGAAACCA	193
Qy	182	CCCAGATTATGGCTTGAATGGAGCTCGCTATCCAAATCCCTCTATTAGGCTACTAA	241
Db	194	CCCAGATTATGGCTTGAATGGAGCTCGCTATCCAAATCCCTCTATTAGGCTACTAA	253
Qy	242	CACCATGACTTTATTAGCTGTCTATCTCGCCCCCTCGCGAGGTTCAATGTTGTTAT	301
Db	254	CACCATGACTTTATTAGCTGTCTATCTCGCCCCCTCGCGAGG-TCAATGTTGTTAT	312
Qy	302	TTCGAAATGCAACAGCTCCGATTAACCCGACATCACTCCAGATGAGGCTTTCTGA	361
Db	313	TTCGAAATGCAACAGCTCCGATTAACCCGACATCACTCCAGATGAGGCTTTCTGA	372
Qy	362	GTGTGGGTCAAATAGTTTCAATGTTCCCAATGGCCCAAACTGACAGTTTAAACGCTG	421
Db	373	GTGTGGGTCAAATAGTTTCAATGTT-CCCAATGGCCCAAACTGACAGTTTAAACGCTG	431
Qy	422	TCTTGGAACTTAATGACAAAGCGGTATCTATCCAGATGACATTAAGTTTGGTTCGT	481
Db	432	TCTTGGAACTTAATGACAAAGCGGTATCTATCCAGATGACATTAAGTTTGGTTCGT	491
Qy	482	TGAATGCTTAACGGCCAGTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC	541
Db	492	TGAATGCTTAACGGCCAGTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC	551
Qy	542	TGTTTGGTATGATGACGAATGCTCAAAAATTAATCTCAATTAAGCTTAGCGGAGTCTC	601
Db	552	TGTTTGGTATGATGACGAATGCTCAAAAATTAATCTCAATTAAGCTTAGCGGAGTCTC	611
Qy	602	TCTATCGCTTCTGAACCCCGGTGACCTGTCGCGAAAGCGAAATGGGGAAC-ACCCGCT	660
Db	612	TCTATCGCTTCTGAACCCCGGTGACCTGTCGCGAAAGCGAAATGGGGAACCAACCCGCT	671
Qy	661	TTTTGGATGATTAATGCAATGT-CTCCAAATGATGCTTCCAAAGATCTCGTGGGAATAC	719
Db	672	TTTTGGATGATTAATGCAATGTCTCCAAATGATGCTTCCAAAGATCTCGTGGGAATAC	731
Qy	720	TGCTGATAGCTTAAGCTTCAATGATCAAAATTAATGCTTAAACCCCTACTTGACA-GCA	778
Db	732	TGCTGATAGCTTAAGCTTCAATGATCAAAATTAATGCTTAAACCCCTACTTGACGCA	791
Qy	779	ATATATAACAGAGGAGCTGCCCTGCTTTTAAACCTTTTTTTTTTATCATCATTTATTAGC	838
Db	792	ATATATAACAGAGGAGCTGCCCTGCTTTTAAACCTTTTTTTTTTATCATCATTTATTAGC	851

Qy	839	TTACTTTCAATAATTGCGACTGGTTCCAAATGACAAGCTTTTGATTTTAACGACTTTTAAC	898
Db	852	TTACTTTCAATAATTGCGACTGGTTCCAAATGACAAGCTTTTGATTTTAACGACTTTTAAC	911
Qy	899	GACAACTTGAAGATCAAAAACAACTAATTTTCGAAGGATCCAAACGATCAGATTTTC	958
Db	912	GACAACTTGAAGATCAAAAACAACTAATTTTCG-----AAACGATCAGATTTTC	963
Qy	959	CTTCAATTTTACTGCGAGTTTATTCGAGCATCTCCGATTTAGCTGCTCCAGTCAACA	1018
Db	964	CATCTATCTTCACTGCTGTTTGTTCGCTGCTTCTGCTTGGCTGCCAGTTAAACA	1023
Qy	1019	CTACAACAGAGATGAACGSCACAAATTCGGCTGAAGCTGTCATCGGTTACTTCAGATT	1078
Db	1024	CCACTACTGAAGAGAGAGAGCTGCTCAAAATTCCTGCTGAGGCTGTCATCGGTTACTCTGACT	1083
Qy	1079	TAGAAGGGGATTTTCGATGTTGCTGTTTTCGCAATTTTCCAAACAGCAAAATAACGGTTAT	1138
Db	1084	TGGAAGGTGACTTCGACGCTGCTGTTTTCGCAATTCCTAACTCTACTAACAACGGTTGT	1143
Qy	1139	TGTTTATAATTAATACTTACTTATTCGAGCATTCGCTGAAGAGAGGGGTATCTCTCGAGA	1198
Db	1144	TGTTTATCAACACTACCATCGCTTCTATCGCTGCTTAAGGAGGAAGGTTTCTCTGGACA	1203
Qy	1199	AAAGAGAGGCTTGAAGCCGAGCCAGGCCAGGCCAGGCCAGGAAAGCAATGTATCA	1258
Db	1204	AGAGAGCTGCTTGTAAAC-----	1220
Qy	1259	TGCCCCCGGGAGAGAGATAAATTTGCTGCTTCCCGGTGTCACGCCAGCAGTGACGG	1318
Db	1221	TGCCAAATCGTCAGAGGTCCATGCTTCTTCCCAAGATGGGCTTTTCGAGC-----	1276
Qy	1319	AGAGAGGTTGCTGTTTTCGATGACAGTGTCCGCGGATTCGCGGTGCTTCCACCCCATGG	1378
Db	1277	-----TGTTAAGGTTAAGTGGCTTTTCCCATACGGTGG	1312
Qy	1379	CCATCGAGAACACTCAAGAGAGAAATGTCCTTCTAACTAGTGGCGGTAGAAATTCCTAG	1438
Db	1313	TTGTCAGAGTAAC--GGTAAAGATTTCTACTCTGAGAAGGAGGTGAGAGAGTACTGTGG	1369
Qy	1439	GGCGCCCGGATTAATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTGGGCACTTA	1498
Db	1370	TGTTCCATAGTAAGAAATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTGGGCA-TTA	1428
Qy	1499	CGAGAAGACCGGCTCTGTAGATTCTAATCAAGAGGATGTCAAGATGCCATTTGCCCTGAG	1558
Db	1429	CGAGAAGACCGGCTCTGTAGATTCTAATCAAGAGGATGTCAAGATGCCATTTGCCCTGAG	1488
Qy	1559	AGATCGAGCTTCAATTTTGTATGATCTTTTATTTGTAACCTATATAGTATAGGATTTTTT	1618
Db	1489	AGATCGAGCTTCAATTTTGTATGATCTTTTATTTGTAACCTATATAGTATAGGATTTTTT	1548
Qy	1619	TTGTCAATTTGTTTCTTCTGCTACGAGCTTGTCTGATCAGCCTATCTCCAGCTGATG	1678
Db	1549	TTGTCAATTTGTTTCTTCTGCTACGAGCTTGTCTGATCAGCCTATCTCCAGCTGATG	1608
Qy	1679	ANATATCTTGTGTAGGGGTTTGGGAAATCAATTCGAGTTTGTATGTTTCTTGATTTTC	1738
Db	1609	ANATATCTTGTGTAGGGGTTTGGGAAATCAATTCGAGTTTGTATGTTTCTTGATTTTC	1668
Qy	1739	CCACTCTCTCCAGAGTACAGAAAGATTAAAGTGAAGTTTCGTTTGTGCAAGCTTATCGAT	1798
Db	1669	CCACTCTCTCCAGAGTACAGAAAGATTAAAGTGAAGTTTCGTTTGTGCAAGCTTATCGAT	1728
Qy	1799	AAGCTTTAATGCGGTAGTTTATCAAGTTAATTCTTAACGAGTCAGGCAACCGTGTATG	1858
Db	1729	AAGCTTTAATGCGGTAGTTTATCAAGTTAATTCTTAACGAGTCAGGCAACCGTGTATG	1789
Qy	1859	AAATCTAACAATGCCCTCATCGTCATCTCGGACCGTCAACCTGATGCTGTAGGCATA	1918
Db	1789	AAATCTAACAATGCCCTCATCGTCATCTCGGACCGTCAACCTGATGCTGTAGGCATA	1848
Qy	1919	GGCTGTTTATGCGGCTACTGCCGGGCTCTTTGCGGGGATATCGTCCATTCGACAGCATC	1978

1849 GGCCTGGTTATGCGGTACTGCGGGCTCTTGGGGATATCGTCCATCCGACAGATC 1908
1979 GCCAGTCACTATGGGTGCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCC 2038
1909 GCCAGTCACTATGGGTGCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCC 1968
2039 GTTCTCGGAGCACTGTCGACCGCTTTGGCGCGCGCCAGCTCGCTCGCTGCTGCTACTT 2098
1969 GTTCTCGGAGCACTGTCGACCGCTTTGGCGCGCGCCAGCTCGCTCGCTGCTGCTACTT 2028
2099 GGAGCCACTATCGACTAGCGATCATGCGGACCAACGCGGCTCTGTGATCTATCGAATC 2158
2029 GGAGCCACTATCGACTAGCGATCATGCGGACCAACGCGGCTCTGTGATCTATCGAATC 2088
2159 TAATGTAAAGTAAATCTTAATAATTAATAAGTCCCAAGTTCTCCATACGAACTT 2218
2089 TAATGTAAAGTAAATCTTAATAATTAATAAGTCCCAAGTTCTCCATACGAACTT 2148
2219 AACAGCATTTGGGTGAGCATCTAGACCTTCAACAGCAGCCAGATCCATCACTGTTGGCC 2278
2149 AACAGCATTTGGGTGAGCATCTAGACCTTCAACAGCAGCCAGATCCATCACTGTTGGCC 2208
2279 AATATGTTTCAGTCCCTCAGAGTTACGCTCTGTGAAAGTGAATCACTGCTGGAGTTGC 2338
2209 AATATGTTTCAGTCCCTCAGAGTTACGCTCTGTGAAAGTGAATCACTGCTGGAGTTGC 2268
2339 AGTGTTAACTCCGCTGATTTGACGGGCATATCCGTAAGTGGCAAGTGGTGGTACC 2398
2268 AGTGTTAACTCCGCTGATTTGACGGGCATATCCGTAAGTGGCAAGTGGTGGTACC 2328
2399 GGAGAGTAAATCTCCAACTCTCTGGAGAGTAGGCAACCAACAGATCCAGCGTG 2458
2329 GGAGAGTAAATCTCCAACTCTCTGGAGAGTAGGCAACCAACAGATCCAGCGTG 2388
2459 TTGTACTTGAATCAATAGAGAGCAATCTCGATTTGAGGATCAAGTCTTCAGGAGC 2518
2389 TTGTACTTGAATCAATAGAGAGCAATCTCGATTTGAGGATCAAGTCTTCAGGAGC 2448
2519 GTACTGATTGGACATTTCCAAAGCCTGCTCGTAGTTGCAACCGATAGGTTGAGAGTG 2578
2449 GTACTGATTGGACATTTCCAAAGCCTGCTCGTAGTTGCAACCGATAGGTTGAGAGTG 2508
2579 TGCAATACACTTGGTACAAATTTCAACCTTTGGCACTGCAAGCTTGGTTGTGAACAGC 2638
2509 TGCAATACACTTGGTACAAATTTCAACCTTTGGCACTGCAAGCTTGGTTGTGAACAGC 2568
2639 ATCTTCAATTTGCGCAAGCTCCTGCTGTATATCGACAGCCCAAGAAATCACTTGGGA 2698
2569 ATCTTCAATTTGCGCAAGCTCCTGCTGTATATCGACAGCCCAAGAAATCACTTGGGA 2628
2699 ATCAATACATGTTGAGTGTGAGACAGAGGTCTGAGGCAACGAAATCTGGATCAGCGTA 2758
2629 ATCAATACATGTTGAGTGTGAGACAGAGGTCTGAGGCAACGAAATCTGGATCAGCGTA 2687
2759 TTTATCAGCAATAAATAGAACTTTAGAAAGGCGCCAGCAGCGATGTCATATCTACACAGGGC 2818
2688 TTTATCAGCAATAAATAGAACTTTAGAAAGGCGCCAGCAGCGATGTCATATCTACACAGGGC 2747
2819 TGATGTGTCATTTGAACCACTTGGCAGCAGTAAACGAACTGGTTTCTGAGCAACAAA 2878
2748 TGATGTGTCATTTGAACCACTTGGCAGCAGTAAACGAACTGGTTTCTGAGCAACAAA 2807
2879 TATTTTGTACACTTAGNACAGTTTCTGTTCCGTAAGCCATAGCAGCTACTGCTGGCC 2938
2808 TATTTTGTACACTTAGNACAGTTTCTGTTCCGTAAGCCATAGCAGCTACTGCTGGCC 2867
2939 GCTCTGCTAGCAGATACACTTAGCACCACTTGTGGGCAACGCTAGATGACTTCTGG 2998
2868 GCTCTGCTAGCAGATACACTTAGCACCACTTGTGGGCAACGCTAGATGACTTCTGG 2927
2999 GGTAAAGGTACCATCTTCTTAGGTGAGATGCAAAAAATTTCTTTGCAACGAGAAC 3058

2928 GGTAAAGGTACCATCTTCTTAGGTGAGATGCAAAAAATTTCTTTGCAACGAGCAAC 2987
3059 TTTGGCAGAAACACCCAGCATCAGGGAAGTGAAGGAGCAAAATTTGGGTTCCACAGGAAT 3118
2988 TTTGGCAGAAACACCCAGCATCAGGGAAGTGAAGGAGCAAAATTTGGGTTCCACAGGAAT 3047
3119 ATAGAGGCCCAACTTTCTCAATAGGTCTTGCAAAAAGAGAGCAGACTACACAGGGCAAGT 3178
3048 ATAGAGGCCCAACTTTCTCAATAGGTCTTGCAAAAAGAGAGCAGACTACACAGGGCAAGT 3107
3179 CTCAACTTTCAGAGCTCCGTTAGTTGAGCTTCATGGATTTCCCTGACGTTATCTATAGA 3238
3108 CTCAACTTTCAGAGCTCCGTTAGTTGAGCTTCATGGAAATTTCCCTGACGTTATCTATAGA 3167
3239 GAGATCAATGGCTCTCTTAACGTTATCTGGCAATTTGCATAAGTTCTCTGGGAAAGGAGC 3298
3168 GAGATCAATGGCTCTCTTAACGTTATCTGGCAATTTGCATAAGTTCTCTGGGAAAGGAGC 3227
3299 TTTCTAACAGAGTGTCTTCAAAAGGAGCTCCATCAAACTTTGGCAGTTAGTTCTTAAAGGGC 3358
3228 TTTCTAACAGAGTGTCTTCAAAAGGAGCTCCATCAAACTTTGGCAGTTAGTTCTTAAAGGGC 3287
3359 TTTGTCCACATTTTGACGAACATTTGTCGACAAATTTGGTTTGAATTTCCATTAATCTGTTTC 3418
3288 TTTGTCCACATTTTGACGAACATTTGTCGACAAATTTGGTTTGAATTTCCATTAATCTGTTTC 3347
3419 CGTTTCTCGATAGGACGAGAGGCAATCTTCAATTTCTGTGAGGAGGCTTGAAGAAC 3478
3348 CGTTTCTCGATAGGACGAGAGGCAATCTTCAATTTCTGTGAGGAGGCTTGAAGAAC 3407
3479 GTCAATTTTGCACAAATTTCAATAGCAGCTTCAAGAGGAGCTTCTTTAGGTTTGGATTTCTTC 3538
3408 GTCAATTTTGCACAAATTTCAATAGCAGCTTCAAGAGGAGCTTCTTTAGGTTTGGATTTCTTC 3467
3539 TTTTAGGTTGTCTCTGCTGTATCTCTGGCTTGGCATCTCTCTTCTTCTAGTGACCTTTAG 3598
3468 TTTTAGGTTGTCTCTGCTGTATCTCTGGCTTGGCATCTCTCTTCTTCTAGTGACCTTTAG 3527
3599 GGAATTCATATCCAGGTTTCTCTCCACCTGCTCCAGCTGTCACACGTCACACGACTTGGCAGATCT 3658
3528 GGAATTCATATCCAGGTTTCTCTCCACCTGCTCCAGCTGTCACACGTCACACGACTTGGCAGATCT 3587
3659 AACTAATGCAAAAATAAATAAGTCAAGCTTCCAGGCTATATCTTCTTCTTGGATTTAGC 3718
3588 AACTAATGCAAAAATAAATAAGTCAAGCTTCCAGGCTATATCTTCTTCTTGGATTTAGC 3647
3719 TTTGCAAGTTTCATAGCTTCTCTCTTAAATTTTAGGCTTCAACAAACTTCTGCTGCTCAAA 3778
3648 TTTGCAAGTTTCATAGCTTCTCTCTTAAATTTTAGGCTTCAACAAACTTCTGCTGCTCAAA 3707
3779 TAAACGTTTGGTATAAGAACCTTCTGAGGATTTGCTTTACGATCCCAAGAGT-GCTTTC 3838
3708 TAAACGTTTGGTATAAGAACCTTCTGAGGATTTGCTTTACGATCCCAAGAGT-GCTTTC 3766
3839 CATGGCTCTAAGACCTTTGATTGGCCAAAACAGAGGAGTGGCTTCCAAAGTGACAGAAACC 3898
3767 CATGGCTCTAAGACCTTTGATTGGCCAAAACAGAGGAGTGGCTTCCAAAGTGACAGAAACC 3826
3899 AACACCTTTGTTTCAACCAAAATTTCAAGAGTCTCCATCAATCCAAATTCGATATCC 3958
3827 AACACCTTTGTTTCAACCAAAATTTCAAGAGTCTCCATCAATCCAAATTCGATATCC 3886
3959 CAGCAACTTTTGAAGTTGCTCCAGATGAGCACCTTTTATACCAAAAACCGGTGACGAGAGA 4018
3887 CAGCAACTTTTGAAGTTGCTCCAGATGAGCACCTTTTATACCAAAAACCGGTGACGAGAGA 3946
4019 TTGGTAGACTCCAGTTTGTGCTTATAGCTCCGGAATAGACTTTTGGACGAGTACAC 4078
3947 TTGGTAGACTCCAGTTTGTGCTTATAGCTCCGGAATAGACTTTTGGACGAGTACAC 4006
4079 CAGGCCAACAGCTAAATAGAGAGTCAAGCAAGTAGTGAATAGACCATCGGGGGC 4138
4007 CAGGCCAACAGCTAAATAGAGAGTCAAGCAAGTAGTGAATAGACCATCGGGGGC 4066

4139 GTCAAGTACCAAGACCCCAACAAATTTCACTGACAGGGAATCTTTTGTGACATCTT CAGA 4198
4067 GTCAAGTACCAAGACCCCAACAAATTTCACTGACAGGGAATCTTTTGTGACATCTT CAGA 4126
4199 AAGTTGATTCAGTAGTCAATTTGCCAGAGATCAATTAATGGGATTTATAGCAAGCAAC 4258
4127 AAGTTGATTCAGTAGTCAATTTGCCAGAGATCAATTAATGGGATTTATAGCAAGCAAC 4186
4259 AGTGGAGTACATCTTCAACCTTTTGGGCTCTCAGAAAAAGCAATTAACAGTTCTACTACC 4318
4187 AGTGGAGTACATCTTCAACCTTTTGGGCTCTCAGAAAAAGCAATTAACAGTTCTACTACC 4246
4319 GCCATTAGTGAATCTTTTCAAAATCGCCAGTGGAGAAAGAAAGGACAGGATCTAGC 4378
4247 GCCATTAGTGAATCTTTTCAAAATCGCCAGTGGAGAAAGAAAGGACAGGATCTAGC 4306
4379 ATTACGGGCAAGATGCACTTTTATCAACAGGGTCTTATAGATTAACCTAGCGCTGG 4438
4307 ATTACGGGCAAGATGCACTTTTATCAACAGGGTCTTATAGATTAACCTAGCGCTGG 4366
4439 GATCATCTTTTGGCAACTTTTCTGCAATCTAGGTCCAAATCACTTCAATGATACC 4498
4367 GATCATCTTTTGGCAACTTTTCTGCAATCTAGGTCCAAATCACTTCAATGATACC 4426
4499 ATTATTGACAACTTGAGCAAGTTGTGATCAGTCTCTCAATTTGGTCTCTCTAAGCGA 4558
4427 ATTATT-----ACGGA 4436
4559 TGACTCAACTTGCACATTAATCTGAAGCTCAGTGGATGAGTGAACCTTATGATAGTTGTG 4618
4437 TGACTCAACTTGCACATTAATCTGAAGCTCAGTGGATGAGTGAACCTTATGATAGTTGTG 4496
4619 CAGCTGTGACAGATAGGAAACACGGCTTTTCTTCAACAACTCAAGAAATTCATA 4678
4497 CAGCTGTGACAGATAGGAAACACGGCTTTTCTTCAACAACTCAAGAAATTCATA 4556
4679 CTCTGCAACTTGGTATGAGTATGAGGAAATGTCTATCTTGAAGTCCGACAGTG 4738
4557 CTCTGCAACTTGGTATGAGTATGAGGAAATGTCTATCTTGAAGTCCGACAGTG 4616
4739 AGTGTAGTCTTGAAGATTTCTGAAGCGGTATTTTATTTATCAGTGTAGTCTATCAGGA 4798
4617 AGTGTAGTCTTGAAGATTTCTGAAGCGGTATTTTATTTATCAGTGTAGTCTATCAGGA 4676
4799 GATCTCTACCGCGAGCATCTGGCGGACCTGACAGTCCGATCACCGGCGCACAGG 4858
4677 GATCTCTACCGCGAGCATCTGGCGGACCTGACAGTCCGATCACCGGCGCACAGG 4724
4859 TCGGTTCTGGCGCTATATTCGCCGATATCCGATGGGAAAGATCGGCTCGCCACTT 4918
4725 TCGGTTCTGGCGCTATATTCGCCGATATCCGATGGGAAAGATCGGCTCGCCACTT 4784
4919 CGGCTCATAGCGCTTTTGGCGGTGGTATGGTGGCGGCGGCTGGCGGCGGACT 4978
4785 CGGCTCATAGCGCTTTTGGCGGTGGTATGGTGGCGGCGGCTGGCGGCGGACT 4844
4979 GTTGGGCGCCATCTCTTGCATGACCACTTCTTGGCGGCGGCTGGCGGCGGCTCAA 5038
4845 GTTGGGCGCCATCTCTTGCATGACCACTTCTTGGCGGCGGCTGGCGGCGGCTCAA 4904
5039 CCTACTAGTGGCTCTCTTAAATGAGAGTCCGATAGGAGAGCGTGGAGTATCTAT 5098
4905 CCTACTAGTGGCTCTCTTAAATGAGAGTCCGATAGGAGAGCGTGGAGTATCTAT 4964
5099 GATTGGAAGTATGGAATGGTATACCGGCACTTCTTCAAGTCTCTTGGGCTCTCTATCAG 5158
4965 GATTGGAAGTATGGAATGGTATACCGGCACTTCTTCAAGTCTCTTGGGCTCTCTATCAG 5024
5159 ATTATGCCCAACTAAGCAACCGGAGGAGGATTTCAATGTAATTTCTGACTTTTG 5218
5025 ATTATGCCCAACTAAGCAACCGGAGGAGGATTTCAATGTAATTTCTGACTTTTG 5084

5219 GTCATCAGTAGACTCGAAGTGTGAGACTATCTCGGTTATGACAGCAGAAATGCTCTTCTT 5278
5085 GTCATCAGTAGACTCGAAGTGTGAGACTATCTCGGTTATGACAGCAGAAATGCTCTTCTT 5144
5279 GGAGACAGTAAATGAAGTCCCAACCAATAAAGAAATCTTGTATCAGGAAACAACTTCTT 5338
5145 GGAGACAGTAAATGAAGTCCCAACCAATAAAGAAATCTTGTATCAGGAAACAACTTCTT 5204
5339 GTTTCGAACTTTTTCGGTCCCTTGAACCTATTAATGATAGATGATATGTCGGGTAGGAA 5398
5205 GTTTCGAACTTTTTCGGTCCCTTGAACCTATTAATGATAGATGATATGTCGGGTAGGAA 5264
5399 TGGAGCGGCAAAATGCTTACCTTCTGGACCTTCAAGAGGTATGTAGGTTTGTAGATACT 5458
5265 TGGAGCGGCAAAATGCTTACCTTCTGGACCTTCAAGAGGTATGTAGGTTTGTAGATACT 5324
5459 GATGCCAACTTTCAGTGAACAGCTTCTTATTTGCTTCAAACTTCCGAAATCCAGAGAAAT 5518
5325 GATGCCAACTTTCAGTGAACAGCTTCTTATTTGCTTCAAACTTCCGAAATCCAGAGAAAT 5384
5519 CAAAGTTCTTGTCTCTATTTGATCCAGCCAGTCCGCTTGTGAACCTGACAAATAGTGTG 5578
5385 CAAAGTTCTTGTCTCTATTTGATCCAGCCAGTCCGCTTGTGAACCTGACAAATAGTGTG 5444
5579 CTCTGCTTTTTCAGGTCATCTTGTATGAATAAATCTAGTCTTGTATCTTAAATAATCTTGA 5638
5445 CTCTGCTTTTTCAGGTCATCTTGTATGAATAAATCTAGTCTTGTATCTTAAATAATCTTGA 5504
5639 CGAGCCAAAGGCGTAATAATACCCAAATCTAAACCTTTTAAAGCTTAAAGGCAAGTA 5698
5505 CGAGCCAAAGGCGTAATAATACCCAAATCTAAACCTTTTAAAGCTTAAAGGCAAGTA 5564
5699 TGCTGCTCTGATTTAAACCCCAATCAGTCTGATCTGATCTCTCATCACTTCAGGCGGC 5758
5585 TGCTGCTCTGATTTAAACCCCAATCAGTCTGATCTGATCTCTCATCACTTCAGGCGGC 5624
5759 ACTATCTTGTTTTGAAGAAATTCGGAGATGCGATATCGAGAAAGAGTACGCTGATTT 5818
5625 ACTATCTTGTTTTGAAGAAATTCGGAGATGCGATATCGAGAAAGAGTACGCTGATTT 5684
5819 TAAACGTGAATTTATCTCAAGATCTCTGCGCTCG 5852
5685 TAAACGTGAATTTATCTCAAGATCGCGCGCGCG 5718

RESULT 2

US-08-358-160-70
; Sequence 70, Application US/08358160
; Patent No. 563143
; GENERAL INFORMATION:
; APPLICANT: LEY, Arthur C.
; APPLICANT: LADNER, Robert C.
; APPLICANT: GUTERMAN, Sonia K.
; APPLICANT: ROBERTS, Bruce L.
; APPLICANT: MARKLAND, William
; APPLICANT: KENT, Rachel B.
; TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ
; TITLE OF INVENTION: DOMAINS THAT INHIBIT HUMAN NEUTROPHIL ELASTASE
; NUMBER OF SEQUENCES: 234
; CORRESPONDENCE ADDRESS:
; ADDRESSES: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W. Suite 300
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/358,160

[illegible]

1619 TTGTCATTTTCTTCTCTGTCAGAGTGTCTCTGATCAGGCTATCTCGAGCTGATG 1678
1549 TTGTCATTTTCTTCTCTGTCAGAGTGTCTCTGATCAGGCTATCTCGAGCTGATG 1608
1679 AATATCTTGTGGTAGGGTTTGGGAAAATCATTCGAGTTTGATGTTTCTTGGTATTC 1738
1609 AATATCTTGTGGTAGGGTTTGGGAAAATCATTCGAGTTTGATGTTTCTTGGTATTC 1668
1739 CCATCTCTCTTCAGAGTACAGAAATTAAGTGAGAGTTCGTTGTGCAAGCTTATCGAT 1798
1669 CCATCTCTCTTCAGAGTACAGAAATTAAGTGAGAGTTCGTTGTGCAAGCTTATCGAT 1728
1799 AAGCTTTAATCGGTAGTTTATACAGATTAATTTGCTAAAGCGAGTACAGGCACTGTATG 1858
1729 AAGCTTTAATCGGTAGTTTATACAGATTAATTTGCTAAAGCGAGTACAGGCACTGTATG 1788
1859 AATCTAACATGCGCTCATCGTCATCTCGGCAACCTGACCCCTGGATGCTGTAGGCATA 1918
1789 AATCTAACATGCGCTCATCGTCATCTCGGCAACCTGACCCCTGGATGCTGTAGGCATA 1848
1919 GCGTTGGTTATGCGGTACTGCGGGGCTCTTGGGGGATATCGTCCATTCGACAGCATC 1978
1849 GCGTTGGTTATGCGGTACTGCGGGGCTCTTGGGGGATATCGTCCATTCGACAGCATC 1908
1979 GCGAGTACTATGCGGTGCTCTAGCGCTATATGCGTTGATGCAATTTCTATCGGCAACC 2038
1909 GCGAGTACTATGCGGTGCTCTAGCGCTATATGCGTTGATGCAATTTCTATCGGCAACC 1968
2039 GTTCTCGGAGCACTGTCTCGACCGCTTTGGCGCGCCCGCAGTCTCTGCTGCTTCTGCTACTT 2098
1969 GTTCTCGGAGCACTGTCTCGACCGCTTTGGCGCGCCCGCAGTCTCTGCTGCTTCTGCTACTT 2028
2099 GGAGCCACTATCGATCGGATCATGGCGAGCAACACCGTCTCTGTTGGATCTATCGAATC 2158
2029 GGAGCCACTATCGATCGGATCATGGCGAGCAACACCGTCTCTGTTGGATCTATCGAATC 2088
2159 TAAATGTAAGTTAAATCTCTAAATATTAATTAAGTCCAGTTTCTCCATACGAACCTT 2218
2089 TAAATGTAAGTTAAATCTCTAAATATTAATTAAGTCCAGTTTCTCCATACGAACCTT 2148
2219 AACAGCATTTGCGGTGAGCATCTAGACCTTCAACAGAGCCAGATCCATCTGCTGGCC 2278
2149 AACAGCATTTGCGGTGAGCATCTAGACCTTCAACAGAGCCAGATCCATCTGCTGGCC 2208
2279 AATATGTTTCACTCCCTCAGGAGTTAGTCTTGTGAGTGTAGTAACTTCTGGAAGTTGC 2338
2209 AATATGTTTCACTCCCTCAGGAGTTAGTCTTGTGAGTGTAGTAACTTCTGGAAGTTGC 2268
2339 AGTGTAACTCCGCTGTATTCAGCGGCATATCCGTTAGTGTGGTAACTGTTGTTGATACC 2398
2269 AGTGTAACTCCGCTGTATTCAGCGGCATATCCGTTAGTGTGGTAACTGTTGTTGATACC 2328
2399 GGAGAGTAATCTCCCAACTCTCTGGAGTAGGCAACCAACAGACAGATCCAGCGTG 2458
2329 GGAGAGTAATCTCCCAACTCTCTGGAGTAGGCAACCAACAGACAGATCCAGCGTG 2388
2459 TTGTACTTGTATCAACATAAGAGAGCAATCTCGATTTGCAAGATCAAGTGTTCAGGAGC 2518
2389 TTGTACTTGTATCAACATAAGAGAGCAATCTCGATTTGCAAGATCAAGTGTTCAGGAGC 2448
2519 GTACTGATTTGACATTTCCAAAGCCTGCTCTAGGTTGCAACCGATAGGTTGTAGAGTG 2578
2449 GTACTGATTTGACATTTCCAAAGCCTGCTCTAGGTTGCAACCGATAGGTTGTAGAGTG 2508
2579 TGCAATACACTTGGCTACAAATTTCAACCCCTTGGCAACTGCAACAGCTTGGTTGTGAACAGC 2638
2509 TGCAATACACTTGGCTACAAATTTCAACCCCTTGGCAACTGCAACAGCTTGGTTGTGAACAGC 2568
2639 ATCTTCAATCTGGCAAGCTCTTGTCTGTGATATCGACAGCCCAACAGATTCACCTGGGA 2698
2569 ATCTTCAATCTGGCAAGCTCTTGTCTGTGATATCGACAGCCCAACAGATTCACCTGGGA 2628

2699 ATCAATACCATGTTTCTGAGTGTGAGACAGAGTCTGAGGCAACGAAATCTGGATCAGGTA 2758
2629 ATCAATACCATGTTTCTGAGTGTGAGACAGAGTCTGAGGCAACGAAATCTGGATCAGGTA 2687
2759 TTTATCAGCAATTAAGTAACTTCAAGAGGCCAGCAGGCAATGTCAATCTACACAGGCG 2818
2688 TTTATCAGCAATTAAGTAACTTCAAGAGGCCAGCAGGCAATGTCAATCTACACAGGCG 2747
2819 TGATGTGCAATTTTGAACCATCATCTTGGAGAGAGTAAGCAACTCTTCTGACCAAAA 2878
2748 TGATGTGCAATTTTGAACCATCATCTTGGAGAGAGTAAGCAACTCTTCTGACCAAAA 2807
2879 TATTTTGTGACAGTATAGGAACAGTTTCTGTTCCGTAAGCCATAGAGCTACTGCTGGGC 2938
2808 TATTTTGTGACAGTATAGGAACAGTTTCTGTTCCGTAAGCCATAGAGCTACTGCTGGGC 2867
2939 GCGTCTCTGCTAGCAGATACACTTATAGCAACCTTGTGGCAACGTAAGTACTTCTGG 2998
2868 GCGTCTCTGCTAGCAGATACACTTATAGCAACCTTGTGGCAACGTAAGTACTTCTGG 2927
2999 GGTAAAGGTAACATCTTCTTGTAGTGTGAGATGCAAAACAAATTTCTTGTCAACAGCAAC 3058
2928 GGTAAAGGTAACATCTTCTTGTAGTGTGAGATGCAAAACAAATTTCTTGTCAACAGCAAC 2987
3059 TTTGCGAGCAACACCCAGCATCAGGGAAGTGAAGGCGAGAAATTCGCGTCCACAGGAAT 3118
2988 TTTGCGAGCAACACCCAGCATCAGGGAAGTGAAGGCGAGAAATTCGCGTCCACAGGAAT 3047
3119 ATAGAGGCAACTTCTCAATAGTCTTGTGAAAACGAGAGCAGACTACACAGGCGAAGT 3178
3048 ATAGAGGCAACTTCTCAATAGTCTTGTGAAAACGAGAGCAGACTACACAGGCGAAGT 3107
3179 CTCAACTTGTCAACGCTCTCGTTAGTGTGAGCTTCTATGGAATTTCTGAGCTTATCTATAGA 3238
3108 CTCAACTTGTCAACGCTCTCGTTAGTGTGAGCTTCTATGGAATTTCTGAGCTTATCTATAGA 3167
3239 GAGATCAATGCTCTCTTAACTTATGCAATTTGCAATGCAATGCTCTGCGGAAGAGC 3298
3168 GAGATCAATGCTCTCTTAACTTATCTGCAATTTGCAATGCAATGCTCTGCGGAAGAGC 3227
3299 TTTTAAACAGAGTGTCTTCAAGGAGCTCCATCAAACTTGGCAGTGTAGTTCTTAAAGGCG 3358
3228 TTTTAAACAGAGTGTCTTCAAGGAGCTCCATCAAACTTGGCAGTGTAGTTCTTAAAGGCG 3287
3359 TTTTGTCAACATTTTGAACCAATTTGCACAATTTGTTTGAATTAATTCATTAATCTGTTTC 3418
3288 TTTTGTCAACATTTTGAACCAATTTGCACAATTTGTTTGAATTAATTCATTAATCTGTTTC 3347
3419 CGTTTCTGATAGGAGCAACGAGGCAATCTCAATTTCTTGTGAGGAGGCTTGTAGAAC 3478
3348 CGTTTCTGATAGGAGCAACGAGGCAATCTCAATTTCTTGTGAGGAGGCTTGTAGAAC 3407
3479 GTCAATTTTGAACCAATTTCAATACGACTTTCAGAGGGAATCTTCTTGTAGGTTTGTTC 3538
3408 GTCAATTTTGAACCAATTTCAATACGACTTTCAGAGGGAATCTTCTTGTAGGTTTGTTC 3467
3539 TTTAGTGTGTTCTCTTGTGATTCCTGGCTTGGCATCTCTTCTCTAGTGAACCTTTAG 3598
3468 TTTAGTGTGTTCTCTTGTGATTCCTGGCTTGGCATCTCTTCTCTAGTGAACCTTTAG 3527
3599 GGACTTCTATCCAGGTTTCTCTCCACCTCGTCCAACTGCAACGCTACACGTAATTTGGCACA 3658
3528 GGACTTCTATCCAGGTTTCTCTCCACCTCGTCCAACTGCAACGCTACACGTAATTTGGCACA 3587
3659 AACTAATGCAAAATTAATAAGTACAGCACTTCCAGGCTATATCTCTTCTTGGATTTAGC 3718
3588 AACTAATGCAAAATTAATAAGTACAGCACTTCCAGGCTATATCTCTTCTTGGATTTAGC 3647
3719 TTTCTCAGTGTATCAGCTTCTCTCTAAATTTTAGCTTCAACAAACTCTCTCTGTCATAA 3778
3648 TTTCTCAGTGTATCAGCTTCTCTCTAAATTTTAGCTTCAACAAACTCTCTCTGTCATAA 3707
3779 TAAACGTTTGGTATTAAGAACTTCTTGGAGCAATTTGCTCTTACGATCCCAACAGGTTGCTTC 3838

3708 TAACCGTTGGTATAGAACCTTTGGAGCATGTCTTACGATCCCAAGGT-GCTTC 3766
3839 CATGCTCTAAGACCTTTGATTGGCCAAAACAGGAAGTGGGTTCCTCAAGTACAGAAACC 3898
3767 CATGCTCTAAGACCTTTGATTGGCCAAAACAGGAAGTGGGTTCCTCAAGTACAGAAACC 3826
3899 AACACCTGTTGTTCAACACCAAAATTCAGAGAGTCTCATACAAATCCAAATTCGATACC 3958
3827 AACACCTGTTGTTCAACACCAAAATTCAGAGAGTCTCATACAAATCCAAATTCGATACC 3886
3959 CAGCAACTTTTCAGTGTCTCAGATGTAGCACTTTTATACCAAAACCGTGCAGCAGAGA 4018
3887 CAGCAACTTTTCAGTGTCTCAGATGTAGCACTTTTATACCAAAACCGTGCAGCAGAGA 3946
4019 TTGAGTACCTCAGTGTGCTTATAGCTTCGGAATAGACTTTTGGACGATACAC 4078
3947 TTGAGTACCTCAGTGTGCTTATAGCTTCGGAATAGACTTTTGGACGATACAC 4006
4079 CAGGCCCAACAGTAAATAGAGAGTACAGCAACCAAGTGTGATACCACTCGGGCG 4138
4007 CAGGCCCAACAGTAAATAGAGAGTACAGCAACCAAGTGTGATACCACTCGGGCG 4066
4139 GTCACTAGTCAAGACGCGCAACAAATTTCACTGACAGGGAACCTTTTGACATCTTCAGA 4198
4067 GTCACTAGTCAAGACGCGCAACAAATTTCACTGACAGGGAACCTTTTGACATCTTCAGA 4126
4199 AGTTTCGTTTTCAGTGTCAATTCGCGAGCATCAATATGGGATATACCAAGCAAC 4258
4127 AGTTTCGTTTTCAGTGTCAATTCGCGAGCATCAATATGGGATATACCAAGCAAC 4186
4259 AGTGAAGTCACTACCACTTTGCGGTCTCAGAAAAGCATAAACAGTTCTACTACC 4318
4187 AGTGAAGTCACTACCACTTTGCGGTCTCAGAAAAGCATAAACAGTTCTACTACC 4246
4319 GCATTAGTGAACCTTTTCAATTCGCCAGTGGAGAGAGAAAGGCAAGCGATCTAGC 4378
4247 GCATTAGTGAACCTTTTCAATTCGCCAGTGGAGAGAGAAAGGCAAGCGATCTAGC 4306
4379 ATTACGGGCAAGGATGCACTTTTATCAACAGAGGTCTCTATAGATAACCTTAGCGCTGG 4438
4307 ATTACGGGCAAGGATGCACTTTTATCAACAGAGGTCTCTATAGATAACCTTAGCGCTGG 4366
4439 GATCATCTTTGGACAACTTTTTCGCAAAATCTAGGTCCTCAAAATCACTTCATTTGATACC 4498
4367 GATCATCTTTGGACAACTTTTTCGCAAAATCTAGGTCCTCAAAATCACTTCATTTGATACC 4426
4499 ATTATGTACAACTTTGACCAAGTGTCTGATCAGCTCTCTCAAAATGGTCTCTGTAACGGA 4558
4427 ATTAT-----ACGGA 4436
4559 TGACTCAACTTGCACATTAACCTTGAAGCTCAGTTCGATTGAGTGAACCTTGATCAGGTTGTG 4618
4437 TGACTCAACTTGCACATTAACCTTGAAGCTCAGTTCGATTGAGTGAACCTTGATCAGGTTGTG 4496
4619 CAGCTGGTCAGCAGATAGGGAACACAGGCTTTTCTTACCAAACTCAAGGAATTTATCAA 4678
4497 CAGCTGGTCAGCAGATAGGGAACACAGGCTTTTCTTACCAAACTCAAGGAATTTATCAA 4556
4679 CTCTGCAACACTTGGTATGACAGGTAGCAAGGGAATGTCTACTTTGAAGTCGACAGTG 4738
4557 CTCTGCAACACTTGGTATGACAGGTAGCAAGGGAATGTCTACTTTGAAGTCGACAGTG 4616
4739 AGTGTAGTCTTGAGAAATTTCTGAAGCGTATTTTATTTATCAGTGTGATCTCATCAGGA 4798
4617 AGTGTAGTCTTGAGAAATTTCTGAAGCGTATTTTATTTATCAGTGTGATCTCATCAGGA 4676
4799 GATCTCTACGCGCGGACGATCGTGGCGGACCTGCGAGTCTCGGATCAACCGGCGCACAGG 4858
4677 GATCTCTACGCGCGGACGATCGTGGC-----CGGATCACCGGCGCACAGG 4724
4859 TCGCGTGTGCGGCTTATCGCGGACATCACCGATCGGGAAGATCGGCTCGCCACTT 4918

4725 TGGGTTGCTGGCGCCTATATCGCGGACATCACCGATGGGGAAGATCGGGCTCGCCACTT 4784
4919 CCGGCTCATGAGCGCTTGTTCGGCGCTGGGTATGTTGGCAGGCCCCGCTGGCCGGGGACT 4978
4785 CCGGCTCATGAGCGCTTGTTCGGCGCTGGGTATGTTGGCAGGCCCCGCTGGCCGGGGACT 4844
4979 GTTGGGCGCCATCTCTTCGATGACCAATCTCTTCGGCGGCGGCTGCTCAACGCGCTCAA 5038
4845 GTTGGGCGCCATCTCTTCGATGACCAATCTCTTCGGCGGCGGCTGCTCAACGCGCTCAA 4904
5039 CTTACTACTGGGCTGCTTCCTAATGAGGAGTCCGATTAAGGGAAGAGCGTCGAGTATCTAT 5098
4905 CTTACTACTGGGCTGCTTCCTAATGAGGAGTCCGATTAAGGGAAGAGCGTCGAGTATCTAT 4964
5099 GATTGGAAGTATGGGAATGATGATACCCGATCTCTTCAGTGTCTTGGAGTCTCTATCAG 5158
4965 GATTGGAAGTATGGGAATGATGATACCCGATCTCTTCAGTGTCTTGGAGTCTCTATCAG 5024
5159 ATTATGCCCAACTTAAAGCAACCGGAGGAGAGATTTTCATGGTAAATTTCTCTGACTTTTG 5218
5025 ATTATGCCCAACTTAAAGCAACCGGAGGAGAGATTTTCATGGTAAATTTCTCTGACTTTTG 5084
5219 GTCATCAGTAGATCGAACTGTGAGACTATCTCGGTTATGACAGCAGAAATGTCCTTCTT 5278
5085 GTCATCAGTAGATCGAACTGTGAGACTATCTCGGTTATGACAGCAGAAATGTCCTTCTT 5144
5279 GGAGACAGTAAATGAAAGTCCCAATTAAGAAATCTCTGTTATCAGGAACAAACTCTT 5338
5145 GGAGACAGTAAATGAAAGTCCCAATTAAGAAATCTCTGTTATCAGGAACAAACTCTT 5204
5339 GTTTT--CGAACTTTTTCGGTGCCTTGAACTATAAAATGTAGAGTGGATATGTCGGTAGG 5396
5205 GTTTTCGGAACCTTTTCGGTGCCTTGAACTATAAAATGTAGAGTGGATATGTCGGTAGG 5264
5397 AATGGAGCGGCAAAATGCTTACCTCTGGACCTTCAAGAGTATGTAGGGTTGTAGATA 5456
5265 AATGGAGCGGCAAAATGCTTACCTCTGGACCTTCAAGAGTATGTAGGGTTGTAGATA 5324
5457 CTGATGCCAACTTCAGTGCACACGTTTCTATTTCGTTCAAACTTCCGAATCCAGAGAA 5516
5325 CTGATGCCAACTTCAGTGCACACGTTTCTATTTCGTTCAAACTTCCGAATCCAGAGAA 5384
5517 ATCAAAGTGTGTTGTCTATTTATGATCCAAAGCAGTGGGCTCTTGAACCTGCAATAGTG 5576
5385 ATCAAAGTGTGTTGTCTATTTATGATCCAAAGCAGTGGGCTCTTGAACCTGCAATAGTG 5444
5577 TGCTCGTGTGTTGAGTGCATCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 5636
5445 TGCTCGTGTGTTGAGTGCATCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 5504
5637 GACGAGCCAGGCGATAAATACCCAAATCTTAAACTCTTTTAAACGTTTAAAGGACAG 5696
5505 GACGAGCCAGGCGATAAATACCCAAATCTTAAACTCTTTTAAACGTTTAAAGGACAG 5564
5697 TATGCTGCTGTGTTTAAACCCCAATCAGCTCGTGTGATGATGATGATGATGATGATGATGAT 5756
5565 TATGCTGCTGTGTTTAAACCCCAATCAGCTCGTGTGATGATGATGATGATGATGATGATGAT 5624
5757 GCATATCTTGTGTTTAAAGAAATTTTCGGAGATGCGATATCGAGAAAAAGGTACGCTGAT 5816
5625 GCATATCTTGTGTTTAAAGAAATTTTCGGAGATGCGATATCGAGAAAAAGGTACGCTGAT 5816
5817 TTTAAACGTAATTTTATCTCAAGATCTCTGCTGCTG 5852
5685 TTTAAACGTAATTTTATCTCAAGATCGCGGCGCG 5720

RESULT 3

US-08-358-160-65
; Sequence 65, Application US/08358160
; Patent No. 5663143
; GENERAL INFORMATION:
; APPLICANT: LEV, Arthur C.

APPLICANT: LADNER, Robert C.
APPLICANT: GUTERMAN, Sonia K.
APPLICANT: ROBERTS, Bruce L.
APPLICANT: MARKLAND, William
APPLICANT: KENT, Rachel B.
TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ
TITLE OF INVENTION: DOMAINS THAT INHIBIT HUMAN NEUTROPHIL ELASTASE
NUMBER OF SEQUENCES: 234
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W. Suite 300
CITY: Washington
STATE: District of Columbia
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/358,160
FILING DATE: 16-DEC-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/133,031
FILING DATE: 13-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,319
FILING DATE: 26-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/664,989
FILING DATE: 01-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/487,063
FILING DATE: 02-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/240,160
FILING DATE: 02-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Cooper, Iver P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: LEV=1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 8157 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: CIRCULAR
MOLECULE TYPE: other nucleic acid
DESCRIPTION: DNA plasmid
US-08-358-160-65

Query Match 55.4%; Score 4565.2; DB 1; Length 8157;
Best Local Similarity 89.9%; Pred. No. 0;
Matches 5262; Conservative 0; Mismatches 11;
Gaps 11;

QY 2 GATCTTAACATCCAAAGACGAAGGTTGAATGAAGGATTTTGGCCATCCGACATCCACAGG 61
DB 14 GATCTTAACATCCAAAGACGAAGGTTGAATGAAGGATTTTGGCCATCCGACATCCACAGG 73
QY 62 TCCATCTTCACACATAAGTGCACAAACGCAACAGGAGGGGATACACTAGCAGCAGACCGTT 121
DB 74 TCCATCTTCACACATAAGTGCACAAACGCAACAGGAGGGGATACACTAGCAGCAGACCGTT 133
QY 122 GCAAAACGAGGAGGACCTCCACTCTCTCTCTCAACACACCCACTTTTGCCATCGAAAAACCA 181
DB 134 GCAAAACGAGGAGGACCTCCACTCTCTCTCTCAACACACCCACTTTTGCCATCGAAAAACCA 193

182 GCCAGTTATTGGGCTTGAATTGAGCTCGCTCAATCCAAATTCCTTCTATTAGGCTACTAA 241
194 GCCAGTTATTGGGCTTGAATTGAGCTCGCTCAATCCAAATTCCTTCTATTAGGCTACTAA 253
242 CACCATGACTTATTAGGCTCTCTATCTGSCCCCTGGGAGGTTTCATGTTGTTTAT 301
254 CACCATGACTTATTAGGCTCTCTATCTGSCCCCTGGGAGG-TCAATGTTGTTTAT 312
302 TTCCGAATGCAACAGCTCCGCATTACACCCGAACATCACTCCAGATGAGGGCTTCTGA 361
313 TTCCGAATGCAACAGCTCCGCATTACACCCGAACATCACTCCAGATGAGGGCTTCTGA 372
362 GTGTGGGTCAAATAGTTTCAATGTTCCCAATGGCCCAAACTGACAGTTTAAACCTG 421
373 GTGTGGGTCAAATAGTTTCAATGTTCCCAATGGCCCAAACTGACAGTTTAAACCTG 431
422 TCTTGGAACTTAATATGACAAAAGCGTGATCTCATCAAGATGAACATAAGTTTGGTTCGT 481
432 TCTTGGAACTTAATATGACAAAAGCGTGATCTCATCAAGATGAACATAAGTTTGGTTCGT 491
482 TGAATGCTAACCGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCCCATACCTTTGTC 541
492 TGAATGCTAACCGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCCCATACCTTTGTC 551
542 TTGTTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAATGCTTAGCGAGTCTC 601
552 TTGTTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAATGCTTAGCGAGTCTC 611
602 TCTATCGCTTCTGAAACCCCGTGCACTGTCGGAAGCAAAATGGGGAAC-ACCCTGCT 660
612 TCTATCGCTTCTGAAACCCCGTGCGACCTGTGCGGAAACGCAAAATGGGGAACCAACCCCTG 671
661 TTTTGGATGATTGATGCTGTCTCCACATGTTGATGCTTCCAGATTTCTGTTGGGATATC 719
672 TTTTGGATGATTGATGCTGTCTCCACATGTTGATGCTTCCAGATTTCTGTTGGGATATC 731
720 TGCTGATAGCTTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTTACTTGACA-GCA 778
732 TGCTGATAGCTTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTTACTTGACAAGCA 791
779 ATATATAACAGAGAGAAAGCTGCTGCTTTAAACCTTTTATCATCATTTATAGC 838
792 ATATATAACAGAGAGAAAGCTGCTGCTTTAAACCTTTTATCATCATTTATAGC 851
839 TTACTTTTCATATTGCGACTGGTTCCAATTCGACAGCTTTTGATTTAACGACTTTTAAC 898
852 TTACTTTTCATATTGCGACTGGTTCCAATTCGACAGCTTTTGATTTAACGACTTTTAAC 911
899 GACAACTTGAGAGATCAAAAACAACTAATTTTCGAAGGATCCAAACGATGAGATTTC 958
912 GACAACTTGAGAGATCAAAAACAACTAATTTTCGAAGGATCCAAACGATGAGATTTC 953
959 CTTCAATTTTACTGAGTTTATTTTCGCGAGCATCTCCGCGATTAGCTGCTCCAGTCAACA 1018
954 ----- 953
1019 CTACACAGAGATGAAACGGCACAATTTCCGGGCTGAAGTGTCTCGTTACTCAGATT 1078
954 ----- 953
1079 TAGAGGGGATTTCCGATGTTGCTGTTTGGCCATTTTCCACAGCACAATAACGGGTTAT 1138
954 ----- 953
1139 TGTTTATTAATACTACTATTGCCAGCAATGCTGTCTAANAAGAGAGGGGTATCTCTCGAGA 1198
954 ----- 953
1199 AAGAGAGGCTGAAGCCAGCCAGGCCAGGCCAGGCCAGGCCAGGAGAGAAACATGTATCA 1258
954 ----- 953
1259 TGGCCCCCGGAGAGGATAAATTTGGCTTCCCGGTTTCCCGGTTTCCCGCCAGCAGTGACGG 1318

Db 954 ----- 953
Qy 1319 AGAGAGTTGCTGTTTGTGATGACAGTGTCCGGGATTCCTGGTGTCTTCCACCCCATGG 1378
Db 954 ----- 953
Qy 1379 CCATCGAGAACACTCAAGNAGNAGTGTCCCTTCTAACTAGTAGGCGGTAGAAATCCCTAG 1438
Db 954 ----- 953
Qy 1439 GCGCGCCGCGAATTAATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTGGGCACTTA 1498
Db 954 -----AGGAATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTGGGCA-TTA 1001
Qy 1499 CGAGAGACCGGTCTGTAGATTCTAATCAAGAGGATGTGAGATGCCATTTGCCTGAG 1558
Db 1002 CGAGAGACCGGTCTGTAGATTCTAATCAAGAGGATGTGAGATGCCATTTGCCTGAG 1061
Qy 1559 AGATCGAGCTTCATTTTGTATCTTTTATTTTGTAACTATATAGTATAGGATTTTTT 1618
Db 1062 AGATCGAGCTTCATTTTGTATCTTTTATTTTGTAACTATATAGTATAGGATTTTTT 1121
Qy 1619 TTGTCAATTTGTTCTTCTGTAGAGCTTGTCTGTATGATCGCTTATCTCGACGTGATG 1678
Db 1122 TTGTCAATTTGTTCTTCTGTAGAGCTTGTCTGTATGATCGCTTATCTCGACGTGATG 1181
Qy 1679 AATATCTTGTGTAGGGTTTGGGAATATCTCGAGTTTGTATGATTTTTTCTTGTATTTT 1738
Db 1182 AATATCTTGTGTAGGGTTTGGGAATATCTCGAGTTTGTATGATTTTTTCTTGTATTTT 1241
Qy 1739 CCATCTCTCTTTCAGAGTACAGAAGATTAAAGTGAAGTTTCTTTGTGCAAGCTTATCGAT 1798
Db 1242 CCATCTCTCTTTCAGAGTACAGAAGATTAAAGTGAAGTTTCTTTGTGCAAGCTTATCGAT 1301
Qy 1799 AAGCTTTAATCGGTAGTTTATACAGTTAAATTCCTAACCGAGTACGACCGGTATG 1858
Db 1302 AAGCTTTAATCGGTAGTTTATACAGTTAAATTCCTAACCGAGTACGACCGGTATG 1361
Qy 1859 AAATCTAAATCGCTCATCGTATCTCGGACCGTCAACCTGGATGCTGTAGGCAATA 1918
Db 1362 AAATCTAAATCGCTCATCGTATCTCGGACCGTCAACCTGGATGCTGTAGGCAATA 1421
Qy 1919 GCTTGTGTTATGCGGTACTGCGGGCTCTTGGGGATATCGTCCATTCGACAGCATC 1978
Db 1422 GCTTGTGTTATGCGGTACTGCGGGCTCTTGGGGATATCGTCCATTCGACAGCATC 1481
Qy 1979 GCCAGTCACTATGGGTGCTGTAGCGCTATATGGTTGATGCAATTTCTATGCGCACCC 2038
Db 1482 GCCAGTCACTATGGGTGCTGTAGCGCTATATGGTTGATGCAATTTCTATGCGCACCC 1541
Qy 2039 GTTCTCGAGCACTGTCCGACCGCTTTGGCGCGCCCGCAGTCTGTGCTTGTGCTACTT 2098
Db 1542 GTTCTCGAGCACTGTCCGACCGCTTTGGCGCGCCCGCAGTCTGTGCTTGTGCTACTT 1601
Qy 2099 GGAGCCACTATCGACTAGCGATCATGGCGACCAACCCGTCCTGTGGATCTATCGAATC 2158
Db 1602 GGAGCCACTATCGACTAGCGATCATGGCGACCAACCCGTCCTGTGGATCTATCGAATC 1661
Qy 2159 TAAATGTAAGTTAAATCTTAAATAATTAATAAGTCCCAAGTTTCTCCATACGAACCTT 2218
Db 1662 TAAATGTAAGTTAAATCTTAAATAATTAATAAGTCCCAAGTTTCTCCATACGAACCTT 1721
Qy 2219 AACAGCATTTGGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCTGCTGGCC 2278
Db 1722 AACAGCATTTGGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCTGCTGGCC 1781
Qy 2279 AATATGTTTCACTCCCTCAGAGTTTACGTCTGTGTGAAGTATGAATCTTGGAGGTTGC 2338
Db 1782 AATATGTTTCACTCCCTCAGAGTTTACGTCTGTGTGAAGTATGAATCTTGGAGGTTGC 1841
Qy 2339 AGTGTAACTCCGCTGTATTCACGGGCATATCCGTAGTTGGCAAGTGTGGTTGATACC 2398
|||||

Db 1842 AGTGTAACTCCGCTGTATTGACGGGCATATCCGTAGCTTGGCAAGTGTGGTTGTATACC 1901
Qy 2399 GGAGGAGTAACTCCCAAACTCTCTGGAGAGTAGGCACCAAAACACAGATCCAGCGTG 2458
Db 1902 GGAGGAGTAACTCCCAAACTCTCTGGAGTAGGCACCAAAACACAGATCCAGCGTG 1961
Qy 2459 TTCTACTTGATCAACATAGAGAGAGCATTTCTGATTTGAGGATCAAGTGTTCAGAGC 2518
Db 1962 TTCTACTTGATCAACATAGAGAGAGCATTTCTGATTTGAGGATCAAGTGTTCAGAGC 2021
Qy 2519 GTACTGATTTGGAATTTCCAAAGCCCTGCTCGTAGTTGCAACCGATAGGTTGTAGAGTG 2578
Db 2022 GTACTGATTTGGAATTTCCAAAGCCCTGCTCGTAGTTGCAACCGATAGGTTGTAGAGTG 2081
Qy 2579 TGAATACACTTCCGTACAAATTTCAACCCCTTGGCAACTGCAACAGCTTGGTTGTGAACAGC 2638
Db 2082 TGAATACACTTTCGCTACAAATTTCAACCCCTTGGCAACTGCAACAGCTTGGTTGTGAACAGC 2141
Qy 2639 ATCTTCAATTTCTGGCAAGCTCTTGTCTGTCATATCGACAGCAACAGAAATCACCTGGGA 2698
Db 2142 ATCTTCAATTTCTGGCAAGCTCTTGTCTGTCATATCGACAGCAACAGAAATCACCTGGGA 2201
Qy 2699 ATCAATACCATTTCAGTTGAGACAGAAAGTCTGAGGCAACGAAATCTGGATCAGCGTA 2758
Db 2202 ATCAATACCATTTCAGTTGAG- CAGAAGTCTGAGGCAACGAAATCTGGATCAGCGTA 2260
Qy 2759 TTTTATCAGCAATAACTAGAACTTCAGAAGCCCGCAGGCGATGCTCAATACTACACAGGC 2818
Db 2261 TTTTATCAGCAATAACTAGAACTTCAGAAGCCCGCAGGCGATGCTCAATACTACACAGGC 2320
Qy 2819 TGATGTGTCAATTTTGAACCATCATCTTGGCAGCAGTAACGAACTGGTTCTCTGGACAAA 2878
Db 2321 TGATGTGTCAATTTTGAACCATCATCTTGGCAGCAGTAACGAACTGGTTCTCTGGACAAA 2380
Qy 2879 TATTTTGTACACTTTAGGAACAGTTTCTGTTCCGTAAAGCCATAGAGCTTACTGCTGGGC 2938
Db 2381 TATTTTGTACACTTTAGGAACAGTTTCTGTTCCGTAAAGCCATAGAGCTTACTGCTGGGC 2440
Qy 2939 GCCTCTGCTAGCACCATACATTAGCAACCACTTGTGGCAACGATAGACTGACTTCTGG 2998
Db 2441 GCCTCTGCTAGCACCATACATTAGCAACCACTTGTGGCAACGATAGACTGACTTCTGG 2500
Qy 2999 GGTAAAGGTACCATCTTCTTAGTGGAGATGCAAAAACAAATTTCTTTCGAACACAGCAAC 3058
Db 2501 GGTAAAGGTACCATCTTCTTAGTGGAGATGCAAAAACAAATTTCTTTCGAACACAGCAAC 2560
Qy 3059 TTTGGCAGAAACACCCAGCATCAGGGAAGTGAAGCAGAAATGGGTTCCACAGGAAT 3118
Db 2561 TTTGGCAGAAACACCCAGCATCAGGGAAGTGAAGCAGAAATGGGTTCCACAGGAAT 2620
Qy 3119 ATAGAGCCCACTTCTCAATAGGTCTTGCAAAACGAGCAGACTTACACAGGCGAAGT 3178
Db 2621 ATAGAGCCCACTTCTCAATAGGTCTTGCAAAACGAGCAGACTTACACAGGCGAAGT 2680
Qy 3179 CTCAACTTGCACAGCTCTCGTTAGTTGAGCTTATCGAATTTCTGACGTTATCTATAGA 3238
Db 2681 CTCAACTTGCACAGCTCTCGTTAGTTGAGCTTATCGAATTTCTGACGTTATCTATAGA 2740
Qy 3239 GAGATCAATGGCTCTCTTAACTTATCTGGCAATTCATAGTTCTCTGGGAAAGGAGC 3298
Db 2741 GAGATCAATGGCTCTCTTAACTTATCTGGCAATTCATAGTTCTCTGGGAAAGGAGC 2800
Qy 3299 TTTCTAACACAGGTGCTTCAAAAGCGACTCCATCAAACTTGGCAGTTAGTTCTTAAAGGCG 3358
Db 2801 TTTCTAACACAGGTGCTTCAAAAGCGACTCCATCAAACTTGGCAGTTAGTTCTTAAAGGCG 2860
Qy 3359 TTTTGTCAACATTTTGAAGCAATTTGTGCAATTTGGTTTGTACTAATTCATATCTGTTTC 3418
Db 2861 TTTTGTCAACATTTTGAAGCAATTTGTGCAATTTGGTTTGTACTAATTCATATCTGTTTC 2920
Qy 3419 CGTTTCTCGATAGGACGAGGAGGCTTCTTCAATTTCTTGTGAGGAGGCTTGAAGAC 3478
Db 2921 CGTTTCTCGATAGGACGAGGAGGCTTCTTCAATTTCTTGTGAGGAGGCTTGAAGAC 2980
|||||

QY 3479 GTCAATTTTGGCAAAATCAATCAAGACCTTTCAGAGGGAATCTTTAGTTGGATCTTC 3338
DB 2981 GTCAATTTTGGCAAAATCAATCAAGACCTTTCAGAGGGAATCTTTAGTTGGATCTTC 3040
QY 3539 TTTAGGTTGTTCTTGGTGTATCTTGGCTTGGCATCTCTCTTCTCTAGTGACCTTTAG 3598
DB 3041 TTTAGGTTGTTCTTGGTGTATCTTGGCTTGGCATCTCTCTTCTCTAGTGACCTTTAG 3100
QY 3599 GGACTTCATATCCAGGTTTCTCTCAACCTGCTTCAACCTGCTCAACCTGCTCAACCTGCTCAACCT 3658
DB 3101 GGACTTCATATCCAGGTTTCTCTCAACCTGCTTCAACCTGCTCAACCTGCTCAACCTGCTCAACCT 3160
QY 3659 AACTAATGCAAAATAAATTAAGTCAGCAATCTCCAGGCTATATCTTCTTGGATTTAGC 3718
DB 3161 AACTAATGCAAAATAAATTAAGTCAGCAATCTCCAGGCTATATCTTCTTGGATTTAGC 3220
QY 3719 TTTCTGCAAGTTCTATCAGCTTCTCTCTTAAATTTTACGTTTCAACAAATCTTCTGCTCAAA 3778
DB 3221 TTTCTGCAAGTTCTATCAGCTTCTCTCTTAAATTTTACGTTTCAACAAATCTTCTGCTCAAA 3280
QY 3779 TAACCGTTTGGTATTAAGAACCTTCTGGAGCATCTCTTACGATCCCAAGGTTGGCTTC 3838
DB 3281 TAACCGTTTGGTATTAAGAACCTTCTGGAGCATCTCTTACGATCCCAAGGTTGGCTTC 3339
QY 3839 CATGGCTCTAAGACCTTGTATGTTGGCCAAATCAGGAAGTGGCTTCCAAAGTGAAGAAACC 3898
DB 3340 CATGGCTCTAAGACCTTGTATGTTGGCCAAATCAGGAAGTGGCTTCCAAAGTGAAGAAACC 3399
QY 3899 AACACCTGTTTGTCTCAACCAACAAATTTCAAGCAGTCTCCATCAAAATCCAAATTCGATACC 3958
DB 3400 AACACCTGTTTGTCTCAACCAACAAATTTCAAGCAGTCTCCATCAAAATTCGATACC 3459
QY 3959 CAGCAAACTTTGAGTTGCTTCAGATGTAGACACCTTTTACCAAAACCGTGAAGACGAGA 4018
DB 3460 CAGCAAACTTTGAGTTGCTTCAGATGTAGACACCTTTTACCAAAACCGTGAAGACGAGA 3519
QY 4019 TTGGTAGACTCCAGTTTGTGCTTATAGCTTCCGGAATAGCTTTTGGACAGATACAC 4078
DB 3520 TTGGTAGACTCCAGTTTGTGCTTATAGCTTCCGGAATAGCTTTTGGACAGATACAC 3579
QY 4079 CAGGCCAAACGAGTAATTAGAAGTGTAGACACCAACCAAGTGTAGTAAGACATCGGGGG 4138
DB 3580 CAGGCCAAACGAGTAATTAGAAGTGTAGACACCAACCAAGTGTAGTAAGACATCGGGGG 3639
QY 4139 GTCAAGTGTCAAGACGCAACAAATTTCACTGACAGGGAATCTTTTGGACATCTTCAGA 4198
DB 3640 GTCAAGTGTCAAGACGCAACAAATTTCACTGACAGGGAATCTTTTGGACATCTTCAGA 3699
QY 4199 AAGTTGCTATTCAGTAGTCAATTTGCGGAGCATCAATTAATGGGGAATATACGAAGCAAC 4258
DB 3700 AAGTTGCTATTCAGTAGTCAATTTGCGGAGCATCAATTAATGGGGAATATATACGAAGCAAC 3759
QY 4259 AGTGGAGTTCATCTACCACTTTGGGCTCTCAGAAAGGCAATTAACAGTTCTACTACC 4318
DB 3760 AGTGGAGTTCATCTACCACTTTGGGCTCTCAGAAAGGCAATTAACAGTTCTACTACC 3819
QY 4319 GCCATTAGTGAATCTTTTCAATTCGCGAGCATCAATTAATGGGGAATATACGAAGCAAC 4378
DB 3820 GCCATTAGTGAATCTTTTCAATTCGCGAGCATCAATTAATGGGGAATATATACGAAGCAAC 3879
QY 4379 ATTACGGGCAAGGATGCACTTTATCAACCAAGGCTCTATAGATAACCTAGCGCTCG 4438
DB 3880 ATTACGGGCAAGGATGCACTTTATCAACCAAGGCTCTATAGATAACCTAGCGCTCG 3939
QY 4439 GATCATCTTTGGCAAACTCTTTCTGGCAATCTAGGTCCAAATCACTTCATTTGATACC 4498
DB 3940 GATCATCTTTGGCAAACTCTTTCTGGCAATCTAGGTCCAAATCACTTCATTTGATACC 3999
QY 4499 ATTAATGTACAACTTGAGCAAGTTGTGATCAGCTCTCAAAATTTGGTCTCTGTAAACGGA 4558
DB 4000 ATTAAT-----ACGGA 4009

QY 4559 TGACATCAACTTGGCAATTAATTAAGCTCAGTCGAATGAGTGAACTTGAATCAGGTTGTG 4618
DB 4010 TGACATCAACTTGGCAATTAATTAAGCTCAGTCGAATGAGTGAACTTGAATCAGGTTGTG 4069
QY 4619 CAGCTGGTTCAGCAGCATAGAGGAAACACGGCTTTTCTTACCAAACTCAAGGAAATATCAAA 4678
DB 4070 CAGCTGGTTCAGCAGCATAGAGGAAACACGGCTTTTCTTACCAAACTCAAGGAAATATCAAA 4129
QY 4679 CTCTGCAAACTTGGCTATGTCAGGTAGCAAGGAAATGTCATATCTTGAAGTCGACAGTG 4738
DB 4130 CTCTGCAAACTTGGCTATGTCAGGTAGCAAGGAAATGTCATATCTTGAAGTCGACAGTG 4189
QY 4739 AGTGTAGTCTTTGAGAAATTTCTGAAGCCGTATTTTATTAATCAAGTGAATCAGTCAATCAGGA 4798
DB 4190 AGTGTAGTCTTTGAGAAATTTCTGAAGCCGTATTTTATTAATCAAGTGAATCAGTCAATCAGGA 4249
QY 4799 GATCCTCTACGCCGGAACGATCTGTCGCGACCTCTGAGGTCTGGATCAACCGGCGCCACAGG 4858
DB 4250 GATCCTCTACGCCGGAACGATCTGTCGCG-----CGGCATCACCGGCGCCACAGG 4297
QY 4859 TGCGGTTGCTGGCCCTTATATCGCCGACATCAACGATGGGGAAGATCGGGCTCGCCACTT 4918
DB 4298 TGCGGTTGCTGGCCCTTATATCGCCGACATCAACGATGGGGAAGATCGGGCTCGCCACTT 4357
QY 4919 CGGGCTCATGAGCGCTTTGTTTTCGCGCGTGGGTATGTTGGCAGGCCCGCTGGCGGGGACT 4978
DB 4358 CGGGCTCATGAGCGCTTTGTTTTCGCGCGTGGGTATGTTGGCAGGCCCGCTGGCGGGGACT 4417
QY 4979 GTTGGGCGCCATCTCTTGCATGACCATCTCTTGGCGCGCGGCTCTCAAACGGCTCAA 5038
DB 4418 GTTGGGCGCCATCTCTTGCATGACCATCTCTTGGCGCGCGGCTCTCAAACGGCTCAA 4477
QY 5039 CCTACTACTGGGCTGCTTCTTAATGAGGAGTGCATTAAGGAGAGCGTCTGAGTATCTAT 5098
DB 4478 CCTACTACTGGGCTGCTTCTTAATGAGGAGTGCATTAAGGAGAGCGTCTGAGTATCTAT 4537
QY 5099 GATTGGAAGTATGGGAATGCTGATACCCGCAATCTTTCAGTGTCTTTCAGGTCTCTCTATCAG 5158
DB 4538 GATTGGAAGTATGGGAATGCTGATACCCGCAATCTTTCAGTGTCTTTCAGGTCTCTCTATCAG 4597
QY 5159 ATTATGCCAACTTAAGCAACCGAGGAGGAGATTTTCATGTTAAATTTCTCTGACTTTTG 5218
DB 4598 ATTATGCCAACTTAAGCAACCGAGGAGGAGATTTTCATGTTAAATTTCTCTGACTTTTG 4657
QY 5219 GTCATCAGTAGACTTCGAACTGTGAGACTATCTCGGTTTATGACAGCAAGAAATGCTCTTCTT 5278
DB 4658 GTCATCAGTAGACTTCGAACTGTGAGACTATCTCGGTTTATGACAGCAAGAAATGCTCTTCTT 4717
QY 5279 GGAGACAGTAAATGAAGTCCCAACAAATAAGAAATCTCTTGTATCAGGAAACAAATCTTCTT 5338
DB 4718 GGAGACAGTAAATGAAGTCCCAACAAATAAGAAATCTCTTGTATCAGGAAACAAATCTTCTT 4777
QY 5339 GTTTCGAACTTTTTCGGTGTCTTGAATATAAGTGTAGAGTGGATATGTCGGGTAGGAA 5398
DB 4778 GTTTCGAACTTTTTCGGTGTCTTGAATATAAGTGTAGAGTGGATATGTCGGGTAGGAA 4837
QY 5399 TGGAGCGGCAAAATGCTTACTCTCGACCTTCAAGAGTATGTAGGTTTGTAGTACT 5458
DB 4838 TGGAGCGGCAAAATGCTTACTCTCGACCTTCAAGAGTATGTAGGTTTGTAGTACT 4897
QY 5459 GATGCCAACTTTCAGTGACACGTTGCTATTTCTTCAAACCAATTCGGAATCCAGAGAAAT 5518
DB 4898 GATGCCAACTTTCAGTGACACGTTGCTATTTCTTCAAACCAATTCGGAATCCAGAGAAAT 4957
QY 5519 CAAAGTTGTTTGTCTACTATTTGATCCAGCGAGTGGGCTTGTAAACCTGAACATAGTGTG 5578
DB 4958 CAAAGTTGTTTGTCTACTATTTGATCCAGCGAGTGGGCTTGTAAACCTGACAAATAGTGTG 5017
QY 5579 CTCTGTGTTTGTAGGTCTATCTTGTATGAATAAATCTAGTCTTTGATCTAAATAAATCTTGA 5638
DB 5018 CTCTGTGTTTGTAGGTCTATCTTGTATGAATAAATCTAGTCTTTGATCTAAATAAATCTTGA 5077
QY 5639 CGAGCGAGCGGAAATTAATACCAATCTTAAACTCTTTTAAACGCTTAAAGCAAGCACTA 5698

Db	2539	ATCACTGCTTGGCCAAATATGTTTTCAGTCCCTCAGGAGTTACGCTTGTGTGAAGTGTGAAC	2480
Qy	2325	TTCTGGAAGGTTGCAGTGTTAACTCCGCTGTATTTGACGGGCATATCCGTACGTTGGCAAA	2384
Db	2479	TTCTGGAAGGTTGCAGTGTTAACTCCGCTGTATTTGACGGGCATATCCGTACGTTGGCAAA	2420
Qy	2385	GTCTGGTTGTATCGGGAGGTAATCTCCACRACCTCTCTGGAGAGTAGGCACCAACAAC	2444
Db	2419	GTCTGGTTGTATCGGGAGGTAATCTCCACRACCTCTCTGGAGAGTAGGCACCAACAAC	2360
Qy	2445	ACAGATCCACGGTGTTGTACTTTCATCAACATAAGAAAGCATTTCTCGATTTTCAGGATC	2504
Db	2359	ACAGATCCACGGTGTTGTACTTTCATCAACATAAGAAAGCATTTCTCGATTTTCAGGATC	2300
Qy	2505	AAGTGTTTCAGAGGCTACTGATTTGGACATTTTCAAAAGCCCTGCTCGTAGGTTGCAACCGAT	2564
Db	2299	AAGTGTTTCAGAGGCTACTGATTTGGACATTTTCAAAAGCCCTGCTCGTAGGTTGCAACCGAT	2240
Qy	2565	AGGGTCTAGAGTGTGCAATACACTTTCGCTGCGTACAAATTTCAACCCCTGGCACTGCACAGCT	2624
Db	2239	AGGGTCTAGAGTGTGCAATACACTTTCGCTGCGTACAAATTTCAACCCCTGGCACTGCACAGCT	2180
Qy	2625	TGGTTGTGAACAGACATCTTCAATCTTGGGAAGTCTCTTGTCTGTCTATATCGACAGCCAAC	2684
Db	2179	TGGTTGTGAACAGACATCTTCAATCTTGGGAAGTCTCTTGTCTGTCTATATCGACAGCCAAC	2120
Qy	2685	AGAAATCACCTGGGAATCAATACCATGTTCACGTTTGACACAGAGGTTCTGAGGCAACGAAA	2744
Db	2119	AGAAATCACCTGGGAATCAATACCATGTTCACGTTTGACACAGAGGTTCTGAGGCAACGAAA	2060
Qy	2745	TCCTGGATCAGCGTATTTTATCAGCAAAATAACTAGAAACTTTCAGAAAGCCCGACAGGCACTGTCA	2804
Db	2059	TCCTGGATCAGCGTATTTTATCAGCAAAATAACTAGAAACTTTCAGAAAGCCCGACAGGCACTGTCA	2000
Qy	2805	ATACTACACAGGCGTGATGTGTCAATTTGAAACCATCATCTTTGGCAGCAGTAACGAACTGG	2864
Db	1999	ATACTACACAGGCGTGATGTGTCAATTTGAAACCATCATCTTTGGCAGCAGTAACGAACTGG	1940
Qy	2865	TTTCTCTGGACCAAAATATTTTGTACACATTTAGGAACAGTTTCTGTTCGTAAAGCCCATAGCA	2924
Db	1939	TTTCTCTGGACCAAAATATTTTGTACACATTTAGGAACAGTTTCTGTTCGTAAAGCCCATAGCA	1880
Qy	2925	GCTACTGCCCTGGGGCCCTCCTGCTAGCACGATACATTTAGCACCAACCTTTGTGGGCAACG	2984
Db	1879	GCTACTGCCCTGGGGCCCTCCTGCTAGCACGATACATTTAGCACCAACCTTTGTGGGCAACG	1820
Qy	2985	TAGATGACTTCTGGGGTAAAGGTAACCATCTTCTTAGTGGAGATGCAAAAACAATTTCT	3044
Db	1819	TAGATGACTTCTGGGGTAAAGGTAACCATCTTCTTAGTGGAGATGCAAAAACAATTTCT	1760
Qy	3045	TTGCAACACAGCAACTTTGGCAGGAACACCCAGCATCAGGGAAGTGGAAAGCAGAAATTCG	3104
Db	1759	TTGCAACACAGCAACTTTGGCAGGAACACCCAGCATCAGGGAAGTGGAAAGCAGAAATTCG	1700
Qy	3105	GTTTCCACCGGAATATAGAGGCCAACTTTTCTCAATAGTCTTCCAAAACGAGAGCAGACT	3164
Db	1699	GTTTCCACCGGAATATAGAGGCCAACTTTTCTCAATAGTCTTCCAAAACGAGAGCAGACT	1640
Qy	3165	ACACCGGGCAAGTCTCAACTTCGCAACAGTCTCCGTTAGTTAGCTTTCATGGAAATTTCCCTG	3224
Db	1639	ACACCGGGCAAGTCTCAACTTCGCAACAGTCTCCGTTAGTTAGCTTTCATGGAAATTTCCCTG	1580
Qy	3225	ACGTTATCTATAGAGAGTCAATGGCTCTCTTAAAGTTATCTGGCAATTCGATAGTTCC	3284
Db	1579	ACGTTATCTATAGAGAGTCAATGGCTCTCTTAAAGTTATCTGGCAATTCGATAGTTCC	1520
Qy	3285	TCTGGGAAAAGAGCTTCTTAACACAGGTGTCTTCAAAAGCGACTCCATCAAACTTTGGCAGTT	3344
Db	1519	TCTGGGAAAAGAGCTTCTTAACACAGGTGTCTTCAAAAGCGACTCCATCAAACTTTGGCAGTT	1460
Qy	3345	AGTTCTAAAAAGGCTTTGTCCACCATTTTGAAGCAACTTTGTGCAAAATTTGTTTGCATTAAT	3404
Db	1459	AGTTCTAAAAAGGCTTTGTCCACCATTTTGAAGCAACTTTGTGCAAAATTTGTTTGCATTAAT	1400

RESULT 4

US-09-029-267-19/c
Sequence 19, Application US/09029267
Patent No. 6107057
GENERAL INFORMATION:
APPLICANT: Crawford, Kenneth
APPLICANT: Zoror, Isabel
APPLICANT: Innis, Michael
TITLE OF INVENTION: Pichia Secretary Leader for Protein
TITLE OF INVENTION: Expression
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: United States
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/029,267
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Chung, Ling-Fong
REGISTRATION NUMBER: 36,482
REFERENCE/DOCKET NUMBER: 1165.100
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2704
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 2659 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-029-267-19

	Query Match	32.3%	Score 2659;	DB 3;	Length 2659;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 2659;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	2145	GGATCTATCGAATCTAAATGTAAGTTAAATCTCTAAATAATTAATAAGTCCACAGTTTC	2204		
Db	2659	GGATCTATCGAATCTAAATGTAAGTTAAATCTCTAAATAATTAATAAGTCCACAGTTTC	2600		
Qy	2205	TCCATACGAACCTTAAACAGCATTCGGGTAGCATCTAGACCTTCAACAGACGACCATCC	2264		
Db	2599	TCCATACGAACCTTAAACAGCATTCGGGTAGCATCTAGACCTTCAACAGACGACCATCC	2540		
Qy	2265	ATCACTGCTTGGCCAAATATGTTTCAGTCCCTCCAGAGTTACGTTCTGTGAAGTGAATGAAC	2324		

3405 TCCATATCTGTTCCGTTTCTGGATAGGACGAGGCGCATCTTCAATTTCTTGTGAG 3464
 Db TCCATATCTGTTCCGTTTCTGGATAGGACGAGGCGCATCTTCAATTTCTTGTGAG 1340
 3465 GAGGCTTTAGAAAGCGTCAATTTTGGCAAAATTAATACGACCTTTCAGAAAGGAGCTTTT 3524
 Db GAGGCTTTAGAAAGCGTCAATTTTGGCAAAATTAATACGACCTTTCAGAAAGGAGCTTTT 1280
 3525 GGTGTTGATTTCTTCTTTAGTTGTTCTTGGTGTATCTTGGCTTGGCATCTCTCTTCCCT 3584
 Db GGTGTTGATTTCTTCTTTAGTTGTTCTTGGTGTATCTTGGCTTGGCATCTCTCTTCCCT 1220
 3585 CTAGTGACCTTTAGGAGCTTCAATCCAGGTTTCTCTCCACCTCGTCCAAAGGTCAACCG 3644
 Db CTAGTGACCTTTAGGAGCTTCAATCCAGGTTTCTCTCCACCTCGTCCAAAGGTCAACCG 1160
 3645 TACTTGGCACTAATCTAATGCAAAATAAATAAGTACGACATCTCCAGGCTATATCT 3704
 Db TACTTGGCACTAATCTAATGCAAAATAAATAAGTACGACATCTCCAGGCTATATCT 1100
 3705 TCCTTGGATTTAGCTTCTGCAAGTTTCATCAGCTTCTCTCCCTTAATTTAGGTTCAACAA 3764
 Db TCCTTGGATTTAGCTTCTGCAAGTTTCATCAGCTTCTCTCCCTTAATTTAGGTTCAACAA 1040
 1099 TCCTTGGATTTAGCTTCTGCAAGTTTCATCAGCTTCTCTCCCTTAATTTAGGTTCAACAA 3824
 Db TCCTTGGATTTAGCTTCTGCAAGTTTCATCAGCTTCTCTCCCTTAATTTAGGTTCAACAA 980
 3825 CACAAGTGGCTTCCATGCTTCAAGCTTCTGATGCTTCAAGCTTCTGATGCTTCAAGCTTCC 3884
 Db CACAAGTGGCTTCCATGCTTCAAGCTTCTGATGCTTCAAGCTTCTGATGCTTCAAGCTTCC 920
 3885 AAGTGACAGAAACCAACACCTGTTGTTGTTCAACACCAATTTCAAGCAGTCTCCATCACA 3944
 Db AAGTGACAGAAACCAACACCTGTTGTTGTTCAACACCAATTTCAAGCAGTCTCCATCACA 860
 3945 TCCAAATTCGATACCCAGCACTTTGATGTTGCTCCAGATAGCAGCTTTTATACCAAA 4004
 Db TCCAAATTCGATACCCAGCACTTTGATGTTGCTCCAGATAGCAGCTTTTATACCAAA 800
 4005 CCGTGACGACGAGATGCTGATCTCCAGTTTGTCTCTTATAGCTTCCGCAATAGACTTT 4064
 Db CCGTGACGACGAGATGCTGATCTCCAGTTTGTCTCTTATAGCTTCCGCAATAGACTTT 740
 4065 TTGACGAGTACACGAGCCCAAGGATTAATAGAGAGTACGACCAACCAAGTAGTGAAT 4124
 Db TTGACGAGTACACGAGCCCAAGGATTAATAGAGAGTACGACCAACCAAGTAGTGAAT 680
 739 TTGACGAGTACACGAGCCCAAGGATTAATAGAGAGTACGACCAACCAAGTAGTGAAT 680
 4125 AGACCATCGGGGGCTGAGTGTACAGAGCCCAACCAATTTCACTGACAGGAACTTT 4184
 Db AGACCATCGGGGGCTGAGTGTACAGAGCCCAACCAATTTCACTGACAGGAACTTT 620
 4185 TTGACATCTTCAGAAAGTTCTGATAGTCAATTTCCGAGCATCAATTAATGGGAT 4244
 Db TTGACATCTTCAGAAAGTTCTGATAGTCAATTTCCGAGCATCAATTAATGGGAT 560
 619 TTGACATCTTCAGAAAGTTCTGATAGTCAATTTCCGAGCATCAATTAATGGGAT 560
 4245 ATACGAGAACCAAGTGGAGTACATCTACCACTTGGGCTTCAGAAAGGATTA 4304
 Db ATACGAGAACCAAGTGGAGTACATCTACCACTTGGGCTTCAGAAAGGATTA 500
 559 ATACGAGAACCAAGTGGAGTACATCTACCACTTGGGCTTCAGAAAGGATTA 500
 4305 ACAGTTCTACTACCGCCCAATTTAGTGAATTTTCAATTCGCCAGTGGAGAAAGGC 4364
 Db ACAGTTCTACTACCGCCCAATTTAGTGAATTTTCAATTCGCCAGTGGAGAAAGGC 440
 4365 ACAGGATCTAGCATTTAGCGGCAAGATGCACTTTTCAACACCGGCTCTATAGATA 4424
 Db ACAGGATCTAGCATTTAGCGGCAAGATGCACTTTTCAACACCGGCTCTATAGATA 380
 4425 ACCCTAGCGCTGGGATCATCTTTTGGCACTCTTTCTGCAATCTAGTCCAAATC 4484
 Db ACCCTAGCGCTGGGATCATCTTTTGGCACTCTTTCTGCAATCTAGTCCAAATC 320

4485 ACTTCATTGATACATTATTGTACAACTTGAGCAAGTTGTCGATCAGCTCCTCAATTTGG 4544
 Db ACTTCATTGATACATTATTGTACAACTTGAGCAAGTTGTCGATCAGCTCCTCAATTTGG 260
 4545 TCCTCTGTACGGATGACTCAACTTGCACATTAACTTGAAGCTCAGTCGATTGAGTGAAC 4604
 Db TCCTCTGTACGGATGACTCAACTTGCACATTAACTTGAAGCTCAGTCGATTGAGTGAAC 200
 4605 TTGATCAGTTTGTGAGCTGCTCAGCAGCATAGGAAACACGGCTTTTCTACAAACTC 4664
 Db TTGATCAGTTTGTGAGCTGCTCAGCAGCATAGGAAACACGGCTTTTCTACAAACTC 140
 4665 AAGGAATTCACAACTTGCACATTGCTGATGTCAGTACGAGGAAATGTCATACTT 4724
 Db AAGGAATTCACAACTTGCACATTGCTGATGTCAGTACGAGGAAATGTCATACTT 80
 4725 GAAGTCGACAGTGTAGTCTTGAATAATCTGAAGCCGTTATTTTATTCAGTGA 4784
 Db GAAGTCGACAGTGTAGTCTTGAATAATCTGAAGCCGTTATTTTATTCAGTGA 20
 4785 CTCAGTCATCAGGAGATCC 4803
 Db CTCAGTCATCAGGAGATCC 1

RESULT 5
 US-08-735-609-4/c
 ; Sequence 4, Application US/08735609
 ; Patent No. 5955360
 ; GENERAL INFORMATION:
 ; APPLICANT: Chamberlain, Jeffrey S.
 ; APPLICANT: Amalfitano, Andrea
 ; APPLICANT: Hauser, Michael A.
 ; APPLICANT: Kumar-Singh, Rajendra
 ; APPLICANT: Hartigan-O'Connor, Dennis J.
 ; TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Medlen & Carroll, LLP
 ; STREET: 220 Montgomery Street, Suite 2200
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: United States Of America
 ; ZIP: 94104
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/735,609
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ingolia, Diane E.
 ; REGISTRATION NUMBER: 40,027
 ; REFERENCE/DOCKET NUMBER: UM-02484
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 705-8410
 ; TELEFAX: (415) 397-8338
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 34303 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: other nucleic acid
 ; DESCRIPTION: /desc = "DNA"
 ; US-08-735-609-4

Query Match 28.0%; Score 2310.8; DB 2; Length 34303;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2356; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

Db 686 ATTAGAAAATAAACAATAGGGTTCGGGCACATTTCCCGAAAAGTCCACCTGAC 627
QY 8065 GTCTAAGAAACCATTTATCATGACATTAACCTATAAAATAGGCGTATCAGGAGCCC 8124
Db 626 GTCTAAGAAACCATTTATCATGACATTAACCTATAAAATAGGCGTATCAGGAGCCC 567
QY 8125 TTTCGCTTCAAGAAATATTCATGATTTGACAGCTTATCATGATAGCT 8176
Db 566 TTTCGCTTCAAG---AATTCATGTTGACAGCTTATCATGATAGCT 519

RESULT 6

US-08-735-609-4/c
; Sequence 4, Application US/08735609
; Patent No. 599432
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; Amalfitano, Andrea
; Hauser, Michael A.
; Kumar-Singh, Rajendra
; Hartigan-O'Connor, Dennis J.
; TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/735, 609
; FILING DATE: 23-Oct-1996
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: UM-02484
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34303 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-735-609-4

Query Match 28.0%; Score 2310.8; DB 2; Length 34303;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2326; Conservative 0; Mismatches 2; Indels 4; Gaps 1;
QY 5845 CTGCTCTCGGGGTTTCGGTGTATGACGTTGAAACCTCTGACATGACAGTCCCGGAGAC 5904
Db 2846 CTGCTCTCGGGGTTTCGGTGTATGACGTTGAAACCTCTGACATGACAGTCCCGGAGAC 2787
QY 5905 GGTACAGCTTGTCTGTAAAGCGGATGCCGGAGCAGACAAGCCCGTCAGGGCCGCTCAGC 5964
Db 2786 GGTACAGCTTGTCTGTAAAGCGGATGCCGGAGCAGACAAGCCCGTCAGGGCCGCTCAGC 2727
QY 5965 GGGTGTTCGGGGTGTCCGGGGCGGACGACATGACCCAGTTCAGTATGAGGAGGTGTA 6024
Db 2726 GGGTGTTCGGGGTGTCCGGGGCGGACGACATGACCCAGTTCAGTATGAGGAGGTGTA 2667

QY 6025 TACTGGCTTAACATATGCGGCATCAGAGCAGATTGTACTGAGAGTGACCATATGCGGTGT 6084
Db 2666 TACTGGCTTAACATATGCGGCATCAGAGCAGATTGTACTGAGAGTGACCATATGCGGTGT 2607
QY 6085 GAAATACCGCAGACATGCGTAAAGGAGAAATACCGATCAGGCGCTCTTCCGCTTCTCG 6144
Db 2606 GAAATACCGCAGACATGCGTAAAGGAGAAATACCGATCAGGCGCTCTTCCGCTTCTCG 2547
QY 6145 CTCAGTACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 6204
Db 2546 CTCAGTACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 2487
QY 6205 CGGGTATACGGTTATCCAGATCAGGGGATAACGAGGAAAGCAATGAGCAAA 6264
Db 2486 CGGGTATACGGTTATCCAGATCAGGGGATAACGAGGAAAGCAATGAGCAAA 2427
QY 6265 GCGCAGCAAAAGGCGCAGAAACCGTAAAGAGCGCGCTTGGTGGCGTTTTCATAGGCTC 6324
Db 2426 GCGCAGCAAAAGGCGCAGAAACCGTAAAGAGCGCGCTTGGTGGCGTTTTCATAGGCTC 2367
QY 6325 CGCCCGCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGCGCAACCGACA 6384
Db 2366 CGCCCGCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGCGCAACCGACA 2307
QY 6385 GAGCTATAAAGATACCGAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTCGTTCGG 6444
Db 2306 GAGCTATAAAGATACCGAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTCGTTCGG 2247
QY 6445 ACCCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAGCGTGGCGCTTCT 6504
Db 2246 ACCCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAGCGTGGCGCTTCT 2187
QY 6505 CATAGCTCAGCGTGTAGTATCTCAGTTCGGTGTAGTTCGCTTCCAGGTGGGCTGT 6564
Db 2186 CATAGCTCAGCGTGTAGTATCTCAGTTCGGTGTAGTTCGCTTCCAGGTGGGCTGT 2127
QY 6565 GTGCAGCAACCCCGCTCAGCCGCGACCGTGGCGCTTATCCGGTAACTATGCTCTGAG 6624
Db 2126 GTGCAGCAACCCCGCTCAGCCGCGACCGTGGCGCTTATCCGGTAACTATGCTCTGAG 2067
QY 6625 TCCAAACCCGGTAAAGACACGACTTATCCCACTGGCAGACCCACTGTAACAGGATTAC 6684
Db 2066 TCCAAACCCGGTAAAGACACGACTTATCCCACTGGCAGACCCACTGTAACAGGATTAC 2007
QY 6685 AGCGGAGGTATGAGCGGTGTACAGAGTCTTGAAGTGGTGGCCTTAACCTACCGCTAC 6744
Db 2006 AGCGGAGGTATGAGCGGTGTACAGAGTCTTGAAGTGGTGGCCTTAACCTACCGCTAC 1947
QY 6745 ACTAAGAGGACAGTATTGTTGTTATCTGCTCTGCTCAAGCCAGTTACCTTCGGAAGAA 6804
Db 1946 ACTAAGAGGACAGTATTGTTGTTATCTGCTCTGCTCAAGCCAGTTACCTTCGGAAGAA 1887
QY 6805 GTTGTAGCTCTTGAATCCGCAAAACCAACCGCTGGTAGCGGTGGTGGTGGTGGTGGT 6864
Db 1886 GTTGTAGCTCTTGAATCCGCAAAACCAACCGCTGGTAGCGGTGGTGGTGGTGGTGGTGGT 1827
QY 6865 AAGCAGCAGATTACCGCGCAAAAAAAGGATCTCAAGAAAGATCTTGAATCTTGTATCTACG 6924
Db 1826 AAGCAGCAGATTACCGCGCAAAAAAAGGATCTCAAGAAAGATCTTGAATCTTGTATCTACG 1767
QY 6925 GGGTCTCAGCTCAGTGGGACGAAACTCAGTAAAGGATTTGGTCAATGAGATTATCA 6984
Db 1766 GGGTCTCAGCTCAGTGGGACGAAACTCAGTAAAGGATTTGGTCAATGAGATTATCA 1707
QY 6985 AAAGAGTCTTCACTAGATCTCTTTTAAATTAAGTAAAGTAAATCAATCTAAAGT 7044
Db 1706 AAAGAGTCTTCACTAGATCTCTTTTAAATTAAGTAAAGTAAATCAATCTAAAGT 1647
QY 7045 ATATATGAGTAAACTTGGTGTGACAGTACCAATGCTTAATCAGTGAGGCACTATCTCA 7104
Db 1646 ATATATGAGTAAACTTGGTGTGACAGTACCAATGCTTAATCAGTGAGGCACTATCTCA 1587

QY 7105 GGGATCTGTCATTTTCGTTTCATCATAGTTCGCTGACATCCCGGTCGTGTAGATAACTACG 7164
Db 1586 GGGATCTGTCATTTTCGTTTCATCATAGTTCGCTGACATCCCGGTCGTGTAGATAACTACG 1527
QY 7165 ATACGGAGGGCTTACCATTCTGGCCCCAGTGTGCAATGATACCGCGAGAGACCCAGCTCA 7224
Db 1526 ATACGGAGGGCTTACCATTCTGGCCCCAGTGTGCAATGATACCGCGAGAGACCCAGCTCA 1467
QY 7225 CCGGCTCCAGATTTATCAGCAATAAACAGCAGCCGGAAGGGCGGAGCGAGAGTGGT 7284
Db 1466 CCGGCTCCAGATTTATCAGCAATAAACAGCAGCCGGAAGGGCGGAGCGAGAGTGGT 1407
QY 7285 CCGTCAACTTTATCCGCTCCATCCAGTCTATTAAATGTTCCCGGGAAGCTAGAGTAAGT 7344
Db 1406 CCGTCAACTTTATCCGCTCCATCCAGTCTATTAAATGTTCCCGGGAAGCTAGAGTAAGT 1347
QY 7345 AGTTCGCCAGTAAATAGTTTGGGMAACGTTGTGTCATGCTGAGGCAATCGTGGTGTCA 7404
Db 1346 AGTTCGCCAGTAAATAGTTTGGGMAACGTTGTGTCATGCTGAGGCAATCGTGGTGTCA 1287
QY 7405 CGCTCGTGGTTGGTATGGCTTCATTAGCTCCGTTCCCAACGATCAAGGCGAGTTACA 7464
Db 1286 CGCTCGTGGTTGGTATGGCTTCATTAGCTCCGTTCCCAACGATCAAGGCGAGTTACA 1227
QY 7465 TGATCCCCCATGTTGTGCAAAAAGCGGTAGCTCTTTCGGTCTCCGATCGTTGTGAGA 7524
Db 1226 TGATCCCCCATGTTGTGCAAAAAGCGGTAGCTCTTTCGGTCTCCGATCGTTGTGAGA 1167
QY 7525 AGTAGTTGGCGGAGTCTTATCACTCATGTTATGGCAGCACTGCATATCTCTTACT 7584
Db 1166 AGTAGTTGGCGGAGTCTTATCACTCATGTTATGGCAGCACTGCATATCTCTTACT 1107
QY 7585 GTCATGCCATCCGTAAGTGTCTTTCTGTGACTGTGTAGTACTCAACCAAGTCAATCTGA 7644
Db 1106 GTCATGCCATCCGTAAGTGTCTTTCTGTGACTGTGTAGTACTCAACCAAGTCAATCTGA 1047
QY 7645 GAATAGTGTATGCGGCGACCGAGTGTCTTTCGCGGGGTCAACACGGGATAATACCGCG 7704
Db 1046 GAATAGTGTATGCGGCGACCGAGTGTCTTTCGCGGGGTCAACACGGGATAATACCGCG 987
QY 7705 CCACATACGAGACTTTTAAAGTGTCTCATCTATGMAAACGTTCTTCGGGCGGAAACTC 7764
Db 986 CCACATACGAGACTTTTAAAGTGTCTCATCTATGMAAACGTTCTTCGGGCGGAAACTC 927
QY 7765 TCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCACTCGTGCACCCAACTGA 7824
Db 926 TCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCACTCGTGCACCCAACTGA 867
QY 7825 TCTTCAGCATCTTTTACTTTTCCAGCGTTTCTCGGTGAGCAAAACAGGAGGCAAAAT 7884
Db 866 TCTTCAGCATCTTTTACTTTTCCAGCGTTTCTCGGTGAGCAAAACAGGAGGCAAAAT 807
QY 7885 GCCGCAAAAAGGGAATAAGGGCGACACGGAATGTTGAATACTCTTCTCTTTT 7944
Db 806 GCCGCAAAAAGGGAATAAGGGCGACACGGAATGTTGAATACTCTTCTCTTTT 747
QY 7945 CAATATTATTGAAGCTTTATCAGGTTTATGTTCTCATGAGCGGATACATATTTGAATGT 8004
Db 746 CAATATTATTGAAGCTTTATCAGGTTTATGTTCTCATGAGCGGATACATATTTGAATGT 687
QY 8005 ATTTAGAAAAATAAACAATAAGGGGTTCGCGGCACTTTCCCGGAAAGTGCCACCTGAC 8064
Db 686 ATTTAGAAAAATAAACAATAAGGGGTTCGCGGCACTTTCCCGGAAAGTGCCACCTGAC 627
QY 8065 GTCTAAGAAACCATTTATCATGACATTAACCTATAAAAAATAGCGGTATCACAGGCC 8124
Db 626 GTCTAAGAAACCATTTATCATGACATTAACCTATAAAAAATAGCGGTATCACAGGCC 567
QY 8125 TTTGCTCTTCAAGAAATTAATCTCATGTTTTCAGAGCTTATCATGATAAGCT 8176
Db 566 TTTGCTCTTCAAG-----AATCTCATGTTTTCAGAGCTTATCATGATAAGCT 519

RESULT 7
US-09-315-372-4/c
; Sequence 4, Application US/09315372
; Patent No. 6057158
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Amalfitano, Andrea
; APPLICANT: Hauser, Michael A.
; APPLICANT: Kumar-Singh, Rajendra
; APPLICANT: Hartigan-O'Connor, Dennis J.
; TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,372
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/735,609
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: UM-02484
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8938
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34303 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
US-09-315-372-4

Query Match 28.0%; Score 2310.8; DB 3; Length 34303;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2326; Conservative 0; Mismatches 2; Indels 4; Gaps 1;
QY 5845 CTGCCTCGCGCGTTTCGGTGATGACGGTGAAAAACCTCTGACACATGCAGCTCCCGGAGAC 5904
Db 2846 CTGCCTCGCGCGTTTCGGTGATGACGGTGAAAAACCTCTGACACATGCAGCTCCCGGAGAC 2787
QY 5905 GGTACAGCTTGTCTGTAAGCGGATGCCCGGAGAGACAAGCCGTCGAGGCGGTCAGC 5964
Db 2786 GGTACAGCTTGTCTGTAAGCGGATGCCCGGAGAGACAAGCCGTCGAGGCGGTCAGC 2727
QY 5965 GGTGTTGGCGGTGTCGGGCGGAGCCATGACCCAGTCACTAGCGATAGCGGAGTGA 6024
Db 2726 GGTGTTGGCGGTGTCGGGCGGAGCCATGACCCAGTCACTAGCGATAGCGGAGTGA 2667
QY 6025 TACTGCTTTAACTATGCGGCATCAGAGCAGATTTGACTGAGAGTGCACCATATCGGTT 6084
Db 2666 TACTGCTTTAACTATGCGGCATCAGAGCAGATTTGACTGAGAGTGCACCATATCGGTT 2607
QY 6085 GAAATACCCGACAGATCGGTAGAGGAAATACCCGATCAGCGGCTCTTCCGCTTCCG 6144
Db 2606 GAAATACCCGACAGATCGGTAGAGGAAATACCCGATCAGCGGCTCTTCCGCTTCCG 2547
QY 6145 CTCACCTGACTCGCTCGGCTCGGTCGTTGCGCTGCGGCGAGCGGTATCAGCTCACTCAAAG 6204

TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wedlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/244,752
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/735,609
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: UM-02484
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 34303 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"
US-09-244-752-4

Query Match 28.0%; Score 2310.8; DB 3; Length 34303;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 2326; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

QY	5845	CTGCCTCCGCGGTTTCGGTGTATGACGGTGAAGAACTCTGACACATGACGATGCCCGGAGAC	5904
DB	2846	CTGCCTCCGCGGTTTCGGTGTATGACGGTGAAGAACTCTGACACATGACGATGCCCGGAGAC	2787
QY	5905	GGTCACAGCTTGTCTGTAAAGCGGATGCCGGAGCAGACAAGCCCGTCAGGCGCGTCAGC	5964
DB	2786	GGTCACAGCTTGTCTGTAAAGCGGATGCCGGAGCAGACAAGCCCGTCAGGCGCGTCAGC	2727
QY	5965	GGGTGTTGGCGGGTTCGGGCGCAGCATGACCCAGTCAGTACGATGACGAGTGTA	6024
DB	2726	GGGTGTTGGCGGGTTCGGGCGCAGCATGACCCAGTCAGTACGATGACGAGTGTA	2667
QY	6025	TACTGGCTTAATATGCGGCATCAGACAGATTGTACTGAGAGTGACCATATGCGGTGT	6084
DB	2666	TACTGGCTTAATATGCGGCATCAGACAGATTGTACTGAGAGTGACCATATGCGGTGT	2607
QY	6085	GAATACCGCACAGATGGTAAAGAGAAATACCGCATCAGCGCTTCGGCTTCCTCG	6144
DB	2606	GAATACCGCACAGATGGTAAAGAGAAATACCGCATCAGCGCTTCGGCTTCCTCG	2547
QY	6145	CTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT	6204
DB	2546	CTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT	2487
QY	6205	CGGTATATACGGTTATCCACAGATCAGGGGATTAACGAGAAAGAAACATGTAGCAAAA	6264
DB	2486	CGGTATATACGGTTATCCACAGATCAGGGGATTAACGAGAAAGAAACATGTAGCAAAA	2427
QY	6265	GGCCAGCAAAAGGCCAGAAACCGTAAAGAAAGCCCGCTTGTGGCTTTTCCATAGGCTC	6324
DB	2426	GGCCAGCAAAAGGCCAGAAACCGTAAAGAAAGCCCGCTTGTGGCTTTTCCATAGGCTC	2367

QY	6325	CGCCCCCTGACGAGCATCAGAAAAATCGACGCTCAAGTCAGAGTGGCGAACCAGCA	6384
DB	2366	CGCCCCCTGACGAGCATCAGAAAAATCGACGCTCAAGTCAGAGTGGCGAACCAGCA	2307
QY	6385	GGACTATAAAGATACCAGGCGTTTCCCTCGTGAAGCTCCCTCGTGGCTCTCTCTGTCG	6444
DB	2306	GGACTATAAAGATACCAGGCGTTTCCCTCGTGAAGCTCCCTCGTGGCTCTCTCTGTCG	2247
QY	6445	ACCTCGCGCTTACCGGATACCTGTCGCGCTTCTCCCTCGGAGAGCTGGGCTTCT	6504
DB	2246	ACCTCGCGCTTACCGGATACCTGTCGCGCTTCTCCCTCGGAGAGCTGGGCTTCT	2187
QY	6505	CATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCTCGCTCCAGCTGGGCTGT	6564
DB	2186	CATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCTCGCTCCAGCTGGGCTGT	2127
QY	6565	GTGCAGAAACCCCGTTTCAGCCGACCCGCTGCGCTTATCCGGTAACTATGCTTTGAG	6624
DB	2126	GTGCAGAAACCCCGTTTCAGCCGACCCGCTGCGCTTATCCGGTAACTATGCTTTGAG	2067
QY	6625	TCCAAACCCGTTAAGACACGACTTATCGCCACTGGGAGAGCCACTGGTAAACAGGATTAC	6684
DB	2066	TCCAAACCCGTTAAGACACGACTTATCGCCACTGGGAGAGCCACTGGTAAACAGGATTAC	2007
QY	6685	AGAGCAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGTGGCTTAACCTACGGCTAC	6744
DB	2006	AGAGCAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGTGGCTTAACCTACGGCTAC	1947
QY	6745	ACTAGAGGACACAGTATTTGGTATCTCGCTCTGCTGACGACGATTAACCTCGGAAAGA	6804
DB	1946	ACTAGAGGACACAGTATTTGGTATCTCGCTCTGCTGACGACGATTAACCTCGGAAAGA	1887
QY	6805	GTGGTAGCTCTTGATCCGCGCAACAAACACCGCTGGTAGCGTGGTTTTTTTGTTCG	6864
DB	1886	GTGGTAGCTCTTGATCCGCGCAACAAACACCGCTGGTAGCGTGGTTTTTTTGTTCG	1827
QY	6865	AGCAGCAGATTACCGCGCAAAAGAAAGATCTCAAGAGATCCTTTGATCTTTTCTAGC	6924
DB	1826	AGCAGCAGATTACCGCGCAAAAGAAAGATCTCAAGAGATCCTTTGATCTTTTCTAGC	1767
QY	6925	GGGTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGATTTTGGTCATGAGATTATCA	6984
DB	1766	GGGTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGATTTTGGTCATGAGATTATCA	1707
QY	6985	AAAGATCTTCACTAGATCTTTTAAATTAAGTATTAAGTATTAAGTATTAAGT	7044
DB	1706	AAAGATCTTCACTAGATCTTTTAAATTAAGTATTAAGTATTAAGTATTAAGT	1647
QY	7045	ATATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGCACCTATCTCA	7104
DB	1646	ATATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGCACCTATCTCA	1587
QY	7105	GGATCTGTATTTCTGTTTCATCTCAGTGTGCTGACTCCCGCTCGTGTAGATACTAGC	7164
DB	1586	GGATCTGTATTTCTGTTTCATCTCAGTGTGCTGACTCCCGCTCGTGTAGATACTAGC	1527
QY	7165	ATACGGGAGGCTTACCATCTGGCCCGAGTCAATGATACCGGAGAGCCAGCTCA	7224
DB	1526	ATACGGGAGGCTTACCATCTGGCCCGAGTCAATGATACCGGAGAGCCAGCTCA	1467
QY	7225	CGGGTCCAGATTTATCAGCAATAAACCCAGCAGCGGAGGAGCGGAGAGAGTGT	7284
DB	1466	CGGGTCCAGATTTATCAGCAATAAACCCAGCAGCGGAGGAGCGGAGAGAGTGT	1407
QY	7285	CTGCAACTTTATCCGCTTCCATCCAGTCTATTAATTTGTCGGGAGAGCTAGAGTAA	7344
DB	1406	CTGCAACTTTATCCGCTTCCATCCAGTCTATTAATTTGTCGGGAGAGCTAGAGTAA	1347
QY	7345	AGTTGCCAGTTAATAGTTTGGCAACCGTTGTCATTTGTCAGGCACTCGTGTGTCA	7404
DB	1346	AGTTGCCAGTTAATAGTTTGGCAACCGTTGTCATTTGTCAGGCACTCGTGTGTCA	1287

```

7405 CGCTCGTGGTGGTATGGCTTCATTCAGCTCCGGTCCCAACGATCAAGGGGAGTTACA 7464
1286 CGCTCGTGGTGGTATGGCTTCATTCAGCTCCGGTCCCAACGATCAAGGGGAGTTACA 1227
7465 TGATCCCGCATGTTGTCACAAAAGCGGTAGCTCTCGGTCCCTCGGATCGTTGTCAGA 7524
1226 TGATCCCGCATGTTGTCACAAAAGCGGTAGCTCTCGGTCCCTCGGATCGTTGTCAGA 1167
7525 AGTAAGTTGGCCGAGGTTATCACTCATGGTTATGGCAGCACTGCATATTTCTTTACT 7584
1166 AGTAAGTTGGCCGAGGTTATCACTCATGGTTATGGCAGCACTGCATATTTCTTTACT 1107
7585 GTCATGCCATCCGTAAGATGTTTCTGTGACCTGGTGAGTACTCAACCAAGTCATTCGA 7644
1106 GTCATGCCATCCGTAAGATGTTTCTGTGACCTGGTGAGTACTCAACCAAGTCATTCGA 1047
7645 GAATAGTGTATGGCGGACCGAGTTGCTCTTGGCCGCGTCAACAGGGGATATACCGGG 7704
1046 GAATAGTGTATGGCGGACCGAGTTGCTCTTGGCCGCGTCAACAGGGGATATACCGGG 987
7705 CCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAAAAAGTTCTTCCGGGGGAAAACTC 7764
986 CCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAAAAAGTTCTTCCGGGGGAAAACTC 927
7765 TCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCACTGTGCAACCACTGA 7824
926 TCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCACTGTGCAACCACTGA 867
7825 TCTTCAGCATCTTTTACTTTTCCACAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAAT 7884
866 TCTTCAGCATCTTTTACTTTTCCACAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAAT 807
7895 CCGCAAAAAGGAATAAGGGGACACAGGAAGTTTGAATATCTATCTTCTTCTTTT 7944
806 CCGCAAAAAGGAATAAGGGGACACAGGAAGTTTGAATATCTATCTTCTTCTTTT 747
7945 CAATATTATCAAGCATTTATCAGGTTTATTTCTCATGAGCGGATACATATTGTAATGT 8004
746 CAATATTATCAAGCATTTATCAGGTTTATTTCTCATGAGCGGATACATATTGTAATGT 687
8005 ATTTAGAAAAATAACAAATAGGGTTCGGGACATTTCCCGAAAAGTCCACCTGAC 8064
686 ATTTAGAAAAATAACAAATAGGGTTCGGGACATTTCCCGAAAAGTCCACCTGAC 627
8065 GTCTAAGAAACCATTTATCATGACATTAACCTTATAAATAAGCGGTATCAGAGGCC 8124
626 GTCTAAGAAACCATTTATCATGACATTAACCTTATAAATAAGCGGTATCAGAGGCC 567
8125 TTTCTGTTTCAAGATTAATTTCTATGTTTTCAGAGTTTATCATCGATAAGCT 8176
566 TTTCTGTTTCAAG-----AATTTCTATGTTTTCAGAGTTTATCATCGATAAGCT 519

```

RESULT 9

US-09-245-497-4/c

; Sequence 4, Application US/09245497

; Patent No. 6083750

; GENERAL INFORMATION:

; APPLICANT: Chamberlain, Jeffrey S.

; APPLICANT: Amalfitano, Andrea

; APPLICANT: Hauser, Michael A.

; APPLICANT: Kumar-Singh, Rajendra

; APPLICANT: Hartigan-O'Connor, Dennis J.

; TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Medlen & Carroll, LLP

; STREET: 220 Montgomery Street, Suite 2200

; CITY: San Francisco

; STATE: California

; COUNTRY: United States Of America

; ZIP: 94104

; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/245,497
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/735,609
; APPLICATION NUMBER: <B> FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: UM-02484
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34303 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
; US-09-245-497-4

Query Match      28.0%; Score 2310.8; DB 3; Length 34303;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2326; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

QY 5845 CTGCTCCGCGGTTTCGGTGTATGACGGTGAACAACTCTGACATGACAGTCCCGGAGAC 5904
DB 2846 CTGCTCCGCGGTTTCGGTGTATGACGGTGAACAACTCTGACATGACAGTCCCGGAGAC 2787
QY 5905 GGTACACAGTTGTCTGTAAAGCGGATGCGGAGGAGACAAAGCCGTCAGGGCGGTCAGC 5964
DB 2786 GGTACACAGTTGTCTGTAAAGCGGATGCGGAGGAGACAAAGCCGTCAGGGCGGTCAGC 2727
QY 5965 GGGTGTTCGGCGGTCGCGGGCGCAGCCATGACCCAGTCACTAGCGGATAGCGAGTGA 6024
DB 2726 GGGTGTTCGGCGGTCGCGGGCGCAGCCATGACCCAGTCACTAGCGGATAGCGAGTGA 2667
QY 6025 TACTGGCTTAATACTAGCGGATCAGAGCAGATTGTACTGAGAGTCCACCATATCGGGTGT 6084
DB 2666 TACTGGCTTAATACTAGCGGATCAGAGCAGATTGTACTGAGAGTCCACCATATCGGGTGT 2607
QY 6085 GAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGGCTCTCCGCTTCCTCG 6144
DB 2606 GAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGGCTCTTCGCTTCCTCG 2547
QY 6145 CTCACCTAGCTCGCTCGGTCGCTCGTTCCGGTTCGGCGGAGCGGTTATCAGTCACTCAAG 6204
DB 2546 CTCACCTAGCTCGCTCGGTCGCTCGTTCCGGTTCGGCGGAGCGGTTATCAGTCACTCAAG 2487
QY 6205 GCGGTAATACGGTTATCCACAGAAATCAGGGGATTAACGAGGAGAAACAAATGTAGCAAAA 6264
DB 2486 GCGGTAATACGGTTATCCACAGAAATCAGGGGATTAACGAGGAGAAACAAATGTAGCAAAA 2427
QY 6265 GGCACGAAAAGGCCAGGAACCGTAAAAAGCGCGGTTGCTGGCGGTTTTCATAGGCTC 6324
DB 2426 GGCACGAAAAGGCCAGGAACCGTAAAAAGCGCGGTTGCTGGCGGTTTTCATAGGCTC 2367
QY 6325 CGCCCCCTGACGAGCATCAGAAAATCGACGCTCAAGTCAAGAGTGCAGAAACCCGACA 6384
DB 2366 CGCCCCCTGACGAGCATCAGAAAATCGACGCTCAAGTCAAGAGTGCAGAAACCCGACA 2307
QY 6385 GGACTATAAGATACCGAGCGGTTTCCCCCTGGAAGCTCCCTCGTGGCTCTCTCTTCG 6444
DB 2306 GGACTATAAGATACCGAGCGGTTTCCCCCTGGAAGCTCCCTCGTGGCTCTCTCTTCG 2247
QY 6445 ACCTGCGCGTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGCAGCTTTCT 6504

```


Db 2246 ACCCTGCCGCTTACCGGATACCTGTCGCCCTTTCTCCCTTCGGGAAGCGTGGCGCTTCT 2187
Qy 6505 CATAGCTCAGCGTGTAGGTATCTCAGTTCGGTGTAGGTCTCGCTCCAAAGCTGGGCTGT 6564
Db 2186 CAACTGCTCAGCGTGTAGGTATCTCAGTTCGGTGTAGGTCTCGCTCCAAAGCTGGGCTGT 2127
Qy 6565 GTGACGAAACCCCGCTTACCGCGACCGCTGGCGCTTATCCGGTAACTATCGCTTTGAG 6624
Db 2126 GTGACGAAACCCCGCTTACCGCGACCGCTGGCGCTTATCCGGTAACTATCGCTTTGAG 2067
Qy 6625 TCCAAACCGGTAAAGACAGACTTATCGCCACTGGCAGCAGCACTGGTAAACAGATTAGC 6684
Db 2066 TCCAAACCGGTAAAGACAGACTTATCGCCACTGGCAGCAGCACTGGTAAACAGATTAGC 2007
Qy 6685 AGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGTCCCTAACTACGGCTAC 6744
Db 2006 AGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGTCCCTAACTACGGCTAC 1947
Qy 6745 ACTAGAGGACAGTATTTGGTATCTGGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAGA 6804
Db 1946 ACTAGAGGACAGTATTTGGTATCTGGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAGA 1887
Qy 6805 GTTGTAGTCTTGTATCCGGCAAAACAAACCCGCTGGTAGCGGTGTTTTTTTGTGTC 6864
Db 1886 GTTGTAGTCTTGTATCCGGCAAAACAAACCCGCTGGTAGCGGTGTTTTTTTGTGTC 1827
Qy 6865 AAGCAGCAGATTACCGCAGAAAAAAGGATCTCAAGAAGTCTCTTTGATCTTTCTACG 6924
Db 1826 AAGCAGCAGATTACCGCAGAAAAAAGGATCTCAAGAAGTCTCTTTGATCTTTCTACG 1767
Qy 6925 GGTCTGACGCTCAGTGTAAACAACTCAGTTAAGGATTTTGGTCTACAGATTATCA 6984
Db 1766 GGTCTGACGCTCAGTGTAAACAACTCAGTTAAGGATTTTGGTCTACAGATTATCA 1707
Qy 6985 AAAAGGATCTTCACTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAAT 7044
Db 1706 AAAAGGATCTTCACTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAAT 1647
Qy 7045 ATATATGATTAACCTTGTGACAGTTACCAATGCTTTAACTAGTGGCAGCTTATCTCA 7104
Db 1646 ATATATGATTAACCTTGTGACAGTTACCAATGCTTTAACTAGTGGCAGCTTATCTCA 1587
Qy 7105 GCGATCTGTATTTCTGTTTCCATCTAGTGTGCTGCTCCCGCTGCTAGATTAACCTACG 7164
Db 1586 GCGATCTGTATTTCTGTTTCCATCTAGTGTGCTGCTCCCGCTGCTAGATTAACCTACG 1527
Qy 7165 ATACGGAGGGCTTACCATCTGGCCCGAGTGTGCAATGATACCGCAGAGCCCGCTCA 7224
Db 1526 ATACGGAGGGCTTACCATCTGGCCCGAGTGTGCAATGATACCGCAGAGCCCGCTCA 1467
Qy 7225 CCGGCTCCAGATTATCAGCAATAAACCCAGCAGCCGGAAGCGCGAGCGAGAGTGGT 7284
Db 1466 CCGGCTCCAGATTATCAGCAATAAACCCAGCAGCCGGAAGCGCGAGCGAGAGTGGT 1407
Qy 7285 CTTGCAACTTTATCCGCTCCATCCAGTCTATTATTTGTCGGGAGCTAGAGTAACT 7344
Db 1406 CTTGCAACTTTATCCGCTCCATCCAGTCTATTATTTGTCGGGAGCTAGAGTAACT 1347
Qy 7345 AGTTCCGCAAGTTAATAGTTTGGCAACGTTGTCATTGTCAGGCACTGCTGGTGTCA 7404
Db 1346 AGTTCCGCAAGTTAATAGTTTGGCAACGTTGTCATTGTCAGGCACTGCTGGTGTCA 1287
Qy 7405 CCGCTCGTGTGTTGGTATGCTTCACTTCCAGCTCCGCTTCCCAACCATCAAGCGAGTTACA 7464
Db 1286 CCGCTCGTGTGTTGGTATGCTTCACTTCCAGCTCCGCTTCCCAACCATCAAGCGAGTTACA 1227
Qy 7465 TGATCCCGCATGTTGTGCAAAAAAGCGGTAGTCTCTTCCGTCCTCCGATGTTGTGAGA 7524
Db 1226 TGATCCCGCATGTTGTGCAAAAAAGCGGTAGTCTCTTCCGTCCTCCGATGTTGTGAGA 1167
Qy 7525 ACTAAGTTGGCGCAGTGTATCACTCATGTTTATGGCAGCACTGCATAATCTCTTACT 7584

Db 1166 AGTAAGTTGGCGCAGTGTATCACTCATGGTTATGGCAGCACTGCATAATTTCTTACT 1107
Qy 7585 GTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTCAATCTGA 7644
Db 1106 GTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTCAATCTGA 1047
Qy 7645 GAATAGTGTATGGCGGACCGAGTGTCTCTTGGCCGCGCTCAACACGGGATATACCGG 7704
Db 1046 GAATAGTGTATGGCGGACCGAGTGTCTCTTGGCCGCGCTCAACACGGGATATACCGG 987
Qy 7705 CCACATAGCAGAACTTTAAAGTGTCTCATCTTGGAAAAACGTTCTTCGGGGCGAAAACTC 7764
Db 986 CCACATAGCAGAACTTTAAAGTGTCTCATCTTGGAAAAACGTTCTTCGGGGCGAAAACTC 927
Qy 7765 TCAAGATCTTACCGCTGTGTAGATCCAGTTCGATGTAAACCCACTGTCGACCCAACTGA 7824
Db 926 TCAAGATCTTACCGCTGTGTAGATCCAGTTCGATGTAAACCCACTGTCGACCCAACTGA 867
Qy 7825 TCTTCAGCATCTTTACTTTTCCACGCTTCTGGGTGAGCAAAAAACAGAGGCAAAAT 7884
Db 866 TCTTCAGCATCTTTACTTTTCCACGCTTCTGGGTGAGCAAAAAACAGAGGCAAAAT 807
Qy 7885 GCCGCAAAAAAGGAAATAAGGGCGACACGAAATGTGAATCTCATCTCTTCTTTT 7944
Db 806 GCCGCAAAAAAGGAAATAAGGGCGACACGAAATGTGAATCTCATCTCTTCTTTT 747
Qy 7945 CAATATATTGAGCACTTTATCAGGCTTATGCTCATGCGGATACATATTTGAATGT 8004
Db 746 CAATATATTGAGCACTTTATCAGGCTTATGCTCATGCGGATACATATTTGAATGT 687
Qy 8005 ATTTAGAAAAATAAACAATAGGGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGAC 8064
Db 686 ATTTAGAAAAATAAACAATAGGGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGAC 627
Qy 8065 GTCTAGAAACCAATTTATCATGACATTAACCTATAAAAAATAGGCGTATCAGAGGCC 8124
Db 626 GTCTAGAAACCAATTTATCATGACATTAACCTATAAAAAATAGGCGTATCAGAGGCC 567
Qy 8125 TTTTCGCTTCAAGAAATTTCTCATGTTTGACAGCTTATCATCGATAAGCT 8176
Db 566 TTTTCGCTTCAAG---AATTCATGTTTGACAGCTTATCATCGATAAGCT 519

RESULT 10

US-09-562-919-4/c

; Sequence 4, Application US/09562919

; Patent No. 6451596

; GENERAL INFORMATION:

; APPLICANT: Chamberlain, Jeffrey S.

; Amalfitano, Andrea

; Hauser, Michael A.

; Kumar-Singh, Rajendra

; Hartigan-O'Connor, Dennis J.

; TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Medlen & Carroll, LLP

; STREET: 220 Montgomery Street, Suite 2200

; CITY: San Francisco

; STATE: California

; COUNTRY: United States Of America

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/562,919

; FILING DATE: 02-May-2000

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/735,609


```
QY 7705 CCACATAGCAGAACTTTAAAGTGGTCTATCTATTCGAAAAAGTCTTCTCGGGCGGAAACTC 7764
Db 986 CCACATAGCAGAACTTTAAAGTGGTCTATCTATTCGAAAAAGTCTTCTCGGGCGGAAACTC 927
QY 7765 TCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAACCAACTGA 7824
Db 926 TCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAACCAACTGA 867
QY 7825 TCTTCAGCATCTTTTACCTTCCACAGCGTTTCTCGGTGAGCAAAAAACAGGAAGCAAAAT 7884
Db 866 TCTTCAGCATCTTTTACCTTCCACAGCGTTTCTCGGTGAGCAAAAAACAGGAAGCAAAAT 807
QY 7885 GCGCCAAAAAGGGAATTAAGGCGGACACGGAATGTTGAATACTCATCTCTTCTCTTTT 7944
Db 806 GCGCCAAAAAGGGAATTAAGGCGGACACGGAATGTTGAATACTCATCTCTTCTCTTTT 747
QY 7945 CAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGACGGGATACATATTTGAATGT 8004
Db 746 CAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGACGGGATACATATTTGAATGT 687
QY 8005 ATTTAGAAAAATAACAAATAGGGTTCCGGCGACATTTCCCGAAAAAGTGCCACCTGAC 8064
Db 686 ATTTAGAAAAATAACAAATAGGGTTCCGGCGACATTTCCCGAAAAAGTGCCACCTGAC 627
QY 8065 GTCTAAGAAACCAATTATCATGACATTAACCTATTAATAATAGGCGGTATCAGGAGGCC 8124
Db 626 GTCTAAGAAACCAATTATCATGACATTAACCTATTAATAATAGGCGGTATCAGGAGGCC 567
QY 8125 TTTCGTCTTAAGAAATTAATTTCTCATGTTTACAGCTTATCATGATAAGCT 8176
Db 566 TTTCGTCTTAAG---AATTCTCATGTTTACAGCTTATCATGATAAGCT 519
```

RESULT 11

```
US-08-125-462-2/c
; Sequence 2, Application US/08125462
; Patent No. 5840840
; GENERAL INFORMATION:
; APPLICANT: Rybak, Susanna M.
; APPLICANT: Youle, Richard J.
; APPLICANT: Newton, Dianne L.
; APPLICANT: Nicholls, Peter J.
; TITLE OF INVENTION: Selective RNase Cytotoxic Reagents
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/125,462
; FILING DATE: 22-SEP-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen L.
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 15280-110-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6727 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..6727
; OTHER INFORMATION: /standard_name="pET-11d-ANG-E6"
US-08-125-462-2
```

```
Query Match 27.9%; Score 2302; DB 2; Length 6727;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2328; Conservative 0; Mismatches 0; Indels 6; Gaps 2;
```

```
QY 5845 CTGCGCTCGCGCGTTTCGGTGATGACGGTCAAAACCTCTCACACATGCGAGTCCCGGAGAC 5904
Db 3821 CTGCGCTCGCGCGTTTCGGTGATGACGGTCAAAACCTCTCACACATGCGAGTCCCGGAGAC 3762
QY 5905 GGTCAACAGCTTGTCTAAGCGGATGCGGGAGCAGACAAGCCCGTCAGGCGCGCTCAGC 5964
Db 3761 GGTCAACAGCTTGTCTAAGCGGATGCGGGAGCAGACAAGCCCGTCAGGCGCGCTCAGC 3702
QY 5965 GGGTGTGGCGGGTGTCCGGGGCGCAGCCATGACCCAGTCACTAGCATAGCGAGTGA 6024
Db 3701 GGGTGTGGCGGGTGTCCGGGGCGCAGCCATGACCCAGTCACTAGCATAGCGAGTGA 3642
QY 6025 TACTGCTTTAACTATATCGGCATCAGAGCAGATTGTACTGAGAGTGCACC--ATATGCGGT 6082
Db 3641 TACTGCTTTAACTATATCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATATGCGGT 3582
QY 6083 GTGAATAACGACACAGATCGTAAGAGAAAAATACCGCATCAGCGCTCTTCGGCTTCCT 6142
Db 3581 GTGAATAACGACACAGATCGTAAGAGAAAAATACCGCATCAGCGCTCTTCGGCTTCCT 3522
QY 6143 CGCTCACTGACTCGCTCGCTCGCTCGCTCGCGGAGCGGTATCAGGTCACTCAA 6202
Db 3521 CGCTCACTGACTCGCTCGCTCGCTCGCTCGCGGAGCGGTATCAGGTCACTCAA 3462
QY 6203 AGCGGTAAATACGTTATCCACAGAAATACGGGATAACCGAGAAAGAACATGTGAGCAA 6262
Db 3461 AGCGGTAAATACGTTATCCACAGAAATACGGGATAACCGAGAAAGAACATGTGAGCAA 3402
QY 6263 AAGGCCAGCAAAAGGCCAGGAAACCGTAAAGGCGCGCTGCTGGCGTTTTTCCATAGGC 6322
Db 3401 AAGGCCAGCAAAAGGCCAGGAAACCGTAAAGGCGCGCTGCTGGCGTTTTTCCATAGGC 3342
QY 6323 TCCGCCCCCTTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGCGGAAACCCGA 6382
Db 3341 TCCGCCCCCTTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGCGGAAACCCGA 3282
QY 6383 CAGGACTATAAGATACAGGCGTTTTCCCGCTGGAGCTCCCTCGTGCGCTCTCCTGTTC 6442
Db 3281 CAGGACTATAAGATACAGGCGTTTTCCCGCTGGAGCTCCCTCGTGCGCTCTCCTGTTC 3222
QY 6443 CGACCTTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTT 6502
Db 3221 CGACCTTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTT 3162
QY 6503 CTCTAGCTCACGCTGTAGGTATCTCAGTTGCGGTAGTGTGTTCTGCTCAAGCTGGGT 6562
Db 3161 CTCTAGCTCACGCTGTAGGTATCTCAGTTGCGGTAGTGTGTTCTGCTCAAGCTGGGT 3102
QY 6563 GTGTGACGAAACCCCGCTTCAGCGCGACCGCTGCGCTTATCCGCTAACTATCGTCTTG 6622
Db 3101 GTGTGACGAAACCCCGCTTCAGCGCGACCGCTGCGCTTATCCGCTAACTATCGTCTTG 3042
QY 6623 AGTCCAAACCGGTGAAGACACGATTATCCCACTGTCAGCAGCCACTGTGTAACAGATTGA 6682
Db 3041 AGTCCAAACCGGTGAAGACACGATTATCCCACTGTCAGCAGCCACTGTGTAACAGATTGA 2982
QY 6683 GCAGAGCGAGGTATGTAGGGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAATACGGCT 6742
Db 2981 GCAGAGCGAGGTATGTAGGGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAATACGGCT 2922
```

6743 ACCTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAA 6802
Db |
2921 ACCTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAA 2862
Qy |
6803 GAGTTGGTAGCTCTTGTATCCGGCAAAACAAACCCAGCTGCTAGCGGTGGTTTTTTTGT 6862
Db |
2861 GAGTTGGTAGCTCTTGTATCCGGCAAAACAAACCCAGCTGCTAGCGGTGGTTTTTTG 2802
Qy |
6863 GCAAGCAGCAGATTACGGCGAGAAAAAGAGATCTCAAGAAGATCTCTTGTATCTTTCTA 6922
Db |
2801 GCAAGCAGCAGATTACGGCGAGAAAAAGAGATCTCAAGAAGATCTCTTGTATCTTTCTA 2742
Qy |
6923 CGGGTCTGAGCTCAGTGGAGCAAAACCTCAAGTTAAGGATTTTGGTCTATCAGATTAT 6982
Db |
2741 CGGGTCTGAGCTCAGTGGAGCAAAACCTCAAGTTAAGGATTTTGGTCTATCAGATTAT 2682
Qy |
6983 CAAAAGGATCTTCACTAGATCTCTTTTAAATTAATAATGAAGTTTAAATCTAATCTAAA 7042
Db |
2681 CAAAAGGATCTTCACTAGATCTCTTTTAAATTAATAATGAAGTTTAAATCTAATCTAAA 2622
Qy |
7043 GTATATAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTAGGCGACCTATCT 7102
Db |
2621 GTATATAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTAGGCGACCTATCT 2562
Qy |
7103 CAGGATCTGTCTATTTTGGTCTATCCATAGTTGCTGACTCCCGTCTGTAGATAACTA 7162
Db |
2561 CAGGATCTGTCTATTTTGGTCTATCCATAGTTGCTGACTCCCGTCTGTAGATAACTA 2502
Qy |
7163 CGATACGGAGGCTTACCTATCGGCCAGTCTGCAATGATACCGGAGCCACGCT 7222
Db |
2501 CGATACGGAGGCTTACCTATCGGCCAGTCTGCAATGATACCGGAGCCACGCT 2442
Qy |
7223 CACGGCTCCAGATTATCAGCAATTAACACGAGCCGAGGCGGAGCGAGAGTG 7282
Db |
2441 CACGGCTCCAGATTATCAGCAATTAACACGAGCCGAGGCGGAGCGAGAGTG 2382
Qy |
7283 GTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGGAAGCTAGAGTAA 7342
Db |
2381 GTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGGAAGCTAGAGTAA 2322
Qy |
7343 GTAGTTCCGCAATTAATAGTTTGGCAACAGTGTGTCATTGTCGAGGATCGTGTGT 7402
Db |
2321 GTAGTTCCGCAATTAATAGTTTGGCAACAGTGTGTCATTGTCGAGGATCGTGTGT 2262
Qy |
7403 CACGCTCGTGTGGTATGCTTCACTCAGCTCCGTTCCCAAGATCAAGCGAGTTA 7462
Db |
2261 CACGCTCGTGTGGTATGCTTCACTCAGCTCCGTTCCCAAGATCAAGCGAGTTA 2202
Qy |
7463 CATGATCCCATGTTGTGCAAAAAGCGGTAGCTCTTCGTCCTCCGATCGTTGCA 7522
Db |
2201 CATGATCCCATGTTGTGCAAAAAGCGGTAGCTCTTCGTCCTCCGATCGTTGCA 2142
Qy |
7523 GAAGTAAGTTGCCCGCAGTTATCACTCATGTTATGGCAGCACTGCATAATTTCTTTA 7582
Db |
2141 GAAGTAAGTTGCCCGCAGTTATCACTCATGTTATGGCAGCACTGCATAATTTCTTTA 2082
Qy |
7583 CTGTATCGCATCCGTATAGATGCTTTTCTGACTGCTGAGTACTCAACCAAGTCATCT 7642
Db |
2081 CTGTATCGCATCCGTATAGATGCTTTTCTGACTGCTGAGTACTCAACCAAGTCATCT 2022
Qy |
7643 GAGATAGTATGGGGGAGCCGAGTTGCTCTTCGCCCGCTCAACAGGGATTAATCCG 7702
Db |
2021 GAGATAGTATGGGGGAGCCGAGTTGCTCTTCGCCCGCTCAACAGGGATTAATCCG 1962
Qy |
7703 CGCCACATAGCAGAACTTTAAAGTGCTCATCTATGGAAACGTTCTTCGGGGGAAAAAC 7762
Db |
1961 CGCCACATAGCAGAACTTTAAAGTGCTCATCTATGGAAACGTTCTTCGGGGGAAAAAC 1902
Qy |
7763 TCTCAAGATCTTACCGTGTGGATGTCAGTGTGATACCACTCGTGACCCCACT 7822
Db |
1901 TCTCAAGATCTTACCGTGTGGATGTCAGTGTGATACCACTCGTGACCCCACT 1842
Qy |
7823 GATCTTCAGCATCTTTTACTTTCACCAAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAA 7882

1841 GATCTTCAGCATCTTTTACTTTTCCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAA 1782
Qy |
7883 ATGCGCGAAAAAAGGAATAGGGCGACACGGAATGTTGAATACTCATACTCTTCCTTT 7942
Db |
1781 ATGCGCGAAAAAAGGAATAGGGCGACACGGAATGTTGAATACTCATACTCTTCCTTT 1722
Qy |
7943 TTCAATATTTTGAAGCATTTTATCAGGGTTATTTCTCATGAGCGGATACATATTTGAAT 8002
Db |
1721 TTCAATATTTTGAAGCATTTTATCAGGGTTATTTCTCATGAGCGGATACATATTTGAAT 1662
Qy |
8003 GTATTGAAAAAATAAACAATAGGGTTCCGCGACATTTCCCGGAAAAAGTGCCACCTG 8062
Db |
1661 GTATTGAAAAAATAAACAATAGGGTTCCGCGACATTTCCCGGAAAAAGTGCCACCTG 1602
Qy |
8063 AGCTCTAAGAACCAATTATTCATGACATTAACCTATAAAATAGGCGTATCAGAGGC 8122
Db |
1601 AGCTCTAAGAACCAATTATTCATGACATTAACCTATAAAATAGGCGTATCAGAGGC 1542
Qy |
8123 CTTTTCTGCTTCAAGAAATAATTTCTCATGTTTGAAGCTTATCATTCGATAAGCT 8176
Db |
1541 CTTTTCTGCTTCAAG---AATTTCTCATGTTTGAAGCTTATCATTCGATAAGCT 1492

RESULT 12

US-08-891-848-2/c
; Sequence 2, Application US/08891848
; Patent No. 5955073
; GENERAL INFORMATION:
; APPLICANT: Rybak, Susanna M.
; APPLICANT: Youle, Richard J.
; APPLICANT: Newton, Dianne L.
; APPLICANT: Nicholls, Peter J.
; TITLE OF INVENTION: Selective Rhase Cytotoxic Reagents
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESS: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,848
; FILING DATE: No. 5955073 yet assigned
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/125,462
; FILING DATE: 22-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/014,082
; FILING DATE: 04-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,195
; FILING DATE: 22-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/510,696
; FILING DATE: 20-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 015280-110310US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6727 base pairs

QY 6863 GCAAGCAGCAGATTACGCGCAGAAAAAGGATCTCAAGAGATCCTTTGATCTTTCTA 6922
Db 2873 GCAAGCAGCAGATTACGCGCAGAAAAAGGATCTCAAGAGATCCTTTGATCTTTCTA 2814
QY 6923 CGGGGTCTGAGCCTCAGTGGACGAAACTCACGTTAAGGGATTTTGGTCATGAGATTAT 6982
Db 2813 CGGGGTCTGAGCCTCAGTGGACGAAACTCACGTTAAGGGATTTTGGTCATGAGATTAT 2754
QY 6983 CAAAAAGGATCTTCAACCTAGATCCTTTTAAATTTAAATGAAGTTTAAATCAATCTAAA 7042
Db 2753 CAAAAAGGATCTTCAACCTAGATCCTTTTAAATTTAAATGAAGTTTAAATCAATCTAAA 2694
QY 7043 GTATATATGAGTAAGTGGTCTGACAGATTACCAATGCTTAAATCAGTGGAGCAGCTATCT 7102
Db 2693 GTATATATGAGTAAGTGGTCTGACAGATTACCAATGCTTAAATCAGTGGAGCAGCTATCT 2634
QY 7103 CAGGATCTGCTATTTCGCTTATCCATAGTTGCTGACTCCCGCTCGTGTAGATACTA 7162
Db 2633 CAGGATCTGCTATTTCGCTTATCCATAGTTGCTGACTCCCGCTCGTGTAGATACTA 2574
QY 7163 CGATACGGAGGGCTTACCAATCTGCGCCAGTGTGCAATGATACCGCAGACCCAGCT 7222
Db 2573 CGATACGGAGGGCTTACCAATCTGCGCCAGTGTGCAATGATACCGCAGACCCAGCT 2514
QY 7223 CACGGCTCCAGATTATCAGCAATAACACGAGCGGAGGCGGAGCGAGAGTG 7282
Db 2513 CACGGCTCCAGATTATCAGCAATAACACGAGCGGAGGCGGAGCGAGAGTG 2454
QY 7283 GTCTGCAACTTTATCCGCTCCATCCAGTCTATTAAATTTGCGGGGAAGCTAGAGTAA 7342
Db 2453 GTCTGCAACTTTATCCGCTCCATCCAGTCTATTAAATTTGCGGGGAAGCTAGAGTAA 2394
QY 7343 GTAGTTCGCCAGTTAATAGTTTGGCAACGTTGTGCAATGCTGCGAGCATCGTGTGT 7402
Db 2393 GTAGTTCGCCAGTTAATAGTTTGGCAACGTTGTGCAATGCTGCGAGCATCGTGTGT 2334
QY 7403 CACGCTCGCTGTTGGTATGCTTCATTCAGTCCGGTCCCAACGATCAAGGGAGTTA 7462
Db 2333 CACGCTCGCTGTTGGTATGCTTCATTCAGTCCGGTCCCAACGATCAAGGGAGTTA 2274
QY 7463 CATGATCCCCCATGTTGCAAAAAAGCGTTAGCTCTCTCGGTCCTCGATGTTGCA 7522
Db 2273 CATGATCCCCCATGTTGCAAAAAAGCGTTAGCTCTCTCGGTCCTCGATGTTGCA 2214
QY 7523 GAAGTAAGTTGGCCGAGTGTACATCATGTTATGGCAGCACTCATTAATTTCTTTA 7582
Db 2213 GAAGTAAGTTGGCCGAGTGTATACATCATGTTATGGCAGCACTCATTAATTTCTTTA 2154
QY 7583 CTGTCATGCCATCCGTAAGATGCTTTTCTGTAAGTGTGAGTACTCAACCAAGTCATTCT 7642
Db 2153 CTGTCATGCCATCCGTAAGATGCTTTTCTGTAAGTGTGAGTACTCAACCAAGTCATTCT 2094
QY 7643 GAGAAATGATATGCGGCGACGAGTTGCTCTTGGCCGCGCTCAACACGGGATATACCG 7702
Db 2093 GAGAAATGATATGCGGCGACGAGTTGCTCTTGGCCGCGCTCAACACGGGATATATACCG 2034
QY 7703 CGCCACATAGCAGAACTTTAAAGTGTCTCATCTTGGAAAAAGCTTTCTCGGGGCGAAAAAC 7762
Db 2033 CGCCACATAGCAGAACTTTAAAGTGTCTCATCTTGGAAAAAGCTTTCTCGGGGCGAAAAAC 1974
QY 7763 TCTCAAGGATCTTACCGTGTGAGATCCAGTTCGATTAACCCACTGTCGACACCCAACT 7822
Db 1973 TCTCAAGGATCTTACCGTGTGAGATCCAGTTCGATTAACCCACTGTCGACACCCAACT 1914
QY 7823 GATCTTCAGCATCTTTTACTTTCACAGCGTTTCTGGGTGAGCAAAAACAGGAGGCAAA 7882
Db 1913 GATCTTCAGCATCTTTTACTTTCACAGCGTTTCTGGGTGAGCAAAAACAGGAGGCAAA 1854
QY 7883 ATGCGCGCAAAAAGGAAATAGGGGCGACACGGAATGTTGAATCTCATCTCTTCCTTT 7942
Db 1853 ATGCGCGCAAAAAGGAAATAGGGGCGACACGGAATGTTGAATCTCATCTCTTCCTTT 1794
QY 7943 TTCAATATATTGAAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTGAA 8002

RESULT 14

US-08-891-848-5/c

; Sequence 5, Application US/08891848

; Patent No. 5955073

; GENERAL INFORMATION:

; APPLICANT: Rybak, Susanna M.

; APPLICANT: Youle, Richard J.

; APPLICANT: Newton, Dianne L.

; APPLICANT: Nicholls, Peter J.

; TITLE OF INVENTION: Selective RNase Cytotoxic Reagents

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/891,848

; FILING DATE: No. 5955073 yet assigned

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/125,462

; FILING DATE: 22-SEP-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/014,082

; FILING DATE: 04-FEB-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/779,195

; FILING DATE: 22-OCT-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/510,696

; FILING DATE: 20-APR-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Weber, Ellen Lauver

; REGISTRATION NUMBER: 32,762

; REFERENCE/DOCKET NUMBER: 015280-110310US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 576-0200

; TELEFAX: (415) 576-0300

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 6799 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA

; FEATURE:

; NAME/KEY: -

; LOCATION: 1..6799

; OTHER INFORMATION: /note= "pET-11d-E6-FB-EDN fusion

OTHER INFORMATION: protein sequence"

US-08-891-848-5

Query Match 27.9%; Score 2302; DB 2; Length 6799;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2328; Conservative 0; Mismatches 0; Indels 6; Gaps 2;

5845 CTGGCTCGCGGTTTCGGTGATGACGGTGAAGAACTCTGACATGACGCTCCGGAGAC 5904
Db CTGGCTCGCGGTTTCGGTGATGACGGTGAAGAACTCTGACATGACGCTCCGGAGAC 3834
5905 GGTACACAGTCTGTGTAAGCGGATCGCGGAGCAGACAGCCCGTCAGGCGCGCTCAGC 5964
Db GGTACACAGTCTGTGTAAGCGGATCGCGGAGCAGACAGCCCGTCAGGCGCGCTCAGC 3774
5965 GGGTGTGGCGGGTGTGGCGGCGCAGCCATGACCCAGTCAAGTACGATAGCGAGTGTA 6024
Db GGGTGTGGCGGGTGTGGCGGCGCAGCCATGACCCAGTCAAGTACGATAGCGAGTGTA 3714
6025 TACTGGCTTAATCTATGGCGCATCAGACAGATTGTACTGAGAGTGACCC--ATATGGGCT 6082
Db TACTGGCTTAATCTATGGCGCATCAGACAGATTGTACTGAGAGTGACCCATATGGGCT 3654
6083 GTGAATAACCCACAGATGCGTAAGGAGAAATACCGCATCAGGCGCTCTTCGGCTTCCT 6142
Db GTGAATAACCCACAGATGCGTAAGGAGAAATACCGCATCAGGCGCTCTTCGGCTTCCT 3594
6143 CGCTCACTGATCGCTGCGCTCGGTGCGGTGCGGCGAGCGGTATCAGTCACTCAA 6202
Db CGCTCACTGATCGCTGCGCTCGGTGCGGCGAGCGGTATCAGTCACTCAA 3534
6203 AGGCGGTAATACGTTATCCACAGAAATCAGGGGTAACGAGGAAAGAAATGTGAGCAA 6262
Db AGGCGGTAATACGTTATCCACAGAAATCAGGGGTAACGAGGAAAGAAATGTGAGCAA 3474
6263 AAGGCCAGCAAAAGCCAGGAAACCGTAAAGAGCGGTTGCTGGCGTTTTCATAGGC 6322
Db AAGGCCAGCAAAAGCCAGGAAACCGTAAAGAGCGGTTGCTGGCGTTTTCATAGGC 3414
6323 TCCGCGCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAAGTGGCGGAAACCCGA 6382
Db TCCGCGCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAAGTGGCGGAAACCCGA 3354
6383 CAGGACTATAAAGATAACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTCTGTT 6442
Db CAGGACTATAAAGATAACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTCTGTT 3294
6443 CGACCTGCGCGTTACCGGATACCTGTCGCGCTTTCCTTCCGGAAGCGTGGCGCTT 6502
Db CGACCTGCGCGTTACCGGATACCTGTCGCGCTTTCCTTCCGGAAGCGTGGCGCTT 3234
6503 CTCTAGCTCAGCGTGTAGTATCTCAGTTCCGGTGTAGGCTCGCTCCAAAGTGGGCT 6562
Db CTCTAGCTCAGCGTGTAGTATCTCAGTTCCGGTGTAGGCTCGCTCCAAAGTGGGCT 3174
6563 GTGTGACGAACCCCGGTTTCAGCGCGAGCGCTGCGCTTATCCGGTAACTATGCTGTTG 6622
Db GTGTGACGAACCCCGGTTTCAGCGCGAGCGCTGCGCTTATCCGGTAACTATGCTGTTG 3114
6623 AGTCCAAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTA 6682
Db AGTCCAAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTA 3054
6683 GCAGAGCGAGTATCTAGGCGGCTGCTACAGAGTTCTTTTGAAGTGGTGGCTTAACTA 6742
Db GCAGAGCGAGTATCTAGGCGGCTGCTACAGAGTTCTTTTGAAGTGGTGGCTTAACTA 2994
6743 ACATAGAGACAGTATTTGGTATCTGCGCTGCTGCTGAGCCAGTTTACCTCGGAAATA 6802
Db ACATAGAGACAGTATTTGGTATCTGCGCTGCTGCTGAGCCAGTTTACCTCGGAAATA 2934
6803 GAGTTGGTAGCTCTTGTATCCCGCAAAACAAACACCGCTGGTGGTGGTGGTGGTGGTGGT 6862

2933 GAGTTGGTAGCTCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTTGGTT 2874
6863 GAACGACGAGATTACGCGCAGAAAAAAGAGATCTCAAGAAAGATCTTTTGAATCTTTCTA 6922
2873 GCAACGACGAGATTACGCGCAGAAAAAAGAGATCTCAAGAAAGATCTTTTGAATCTTTCTA 2814
6923 CGGGTCTGACGCTCAGTGGAAAGAACTCAGCTTAAAGGATTTGGTCAAGAGATTAT 6982
2813 CGGGTCTGACGCTCAGTGGAAAGAACTCAGCTTAAAGGATTTGGTCAAGAGATTAT 2754
6983 CAAAAGGATCTTCACTTACCTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 7042
2753 CAAAAGGATCTTCACTTACCTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 2694
7043 GTATATATGAGTAACCTTGGTCTGACGTTACCAATGCTTAAATCAGTGAGGACCTATCT 7102
2693 GTATATATGAGTAACCTTGGTCTGACGTTACCAATGCTTAAATCAGTGAGGACCTATCT 2634
7103 CAGCGATCTGCTATTTCTGTTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 7162
2633 CAGCGATCTGCTATTTCTGTTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 2574
7163 CGATACGGAGGGCTTACCATCTGCGCCAGTGTGCAATGATACCGGAGACCCAGCT 7222
2573 CGATACGGAGGGCTTACCATCTGCGCCAGTGTGCAATGATACCGGAGACCCAGCT 2514
7223 CACCGCTCCAGATTTATCAGCAATTAACAGCAGCCGGAAGGCGCAGCGCAGAAAGTG 7282
2513 CACCGCTCCAGATTTATCAGCAATTAACAGCAGCCGGAAGGCGCAGCGCAGAAAGTG 2454
7283 GTCTGCAACTTTATCCGCTTCCATCCAGTCTATTAATTTGTCGGGAAAGCTAGAGTAA 7342
2453 GTCTGCAACTTTATCCGCTTCCATCCAGTCTATTAATTTGTCGGGAAAGCTAGAGTAA 2394
7343 GTAGTTCCGAGTTAATAGTTTGGCAACGTTGTTGCCATGCTGCGAGGATCTGGTGT 7402
2393 GTAGTTCCGAGTTAATAGTTTGGCAACGTTGTTGCCATGCTGCGAGGATCTGGTGT 2334
7403 CACGCTCGCTGCTTGGTATGCTTCACTTCACTTCCGCTTCCCAACGATCAAGCGAGTTA 7462
2333 CACGCTCGCTGCTTGGTATGCTTCACTTCACTTCCGCTTCCCAACGATCAAGCGAGTTA 2274
7463 CATGATCCCGCATGTTGTGCAAAAAGCGTTAGCTTCTCGTCCCTCGATCTGTTGCA 7522
2273 CATGATCCCGCATGTTGTGCAAAAAGCGTTAGCTTCTCGTCCCTCGATCTGTTGCA 2214
7523 GAAGTAAGTTGCGCGAGTTTATCAGTCACTGTTTGGCAGCACTGCATATTTCTCTTA 7582
2213 GAAGTAAGTTGCGCGAGTTTATCAGTCACTGTTTGGCAGCACTGCATATTTCTCTTA 2154
7583 CTGTCAATCCCATCCGTAAGATCTTTTCTGATCTGATGAGTACTCAACCAAGTCACTCT 7642
2153 CTGTCAATCCCATCCGTAAGATCTTTTCTGATCTGATGAGTACTCAACCAAGTCACTCT 2094
7643 GAGATAGTGTATCGCGCAGCGAGTTGCTCTTCCCGCGCTCAACAGGAGTAAATACCG 7702
2093 GAGATAGTGTATCGCGCAGCGAGTTGCTCTTCCCGCGCTCAACAGGAGTAAATACCG 2034
7703 CGCCACATAGCAGAACTTTTAAAGTGTCTCATCTTGGAAAAAGTCTTTCGGGGCGAAAAAC 7762
2033 CGCCACATAGCAGAACTTTTAAAGTGTCTCATCTTGGAAAAAGTCTTTCGGGGCGAAAAAC 1974
7763 TCTCAAGATCTTACCGCTGTTGAGTCCAGTTCGATGTAAACCCCACTCTGTCACCCCACT 7822
1973 TCTCAAGATCTTACCGCTGTTGAGTCCAGTTCGATGTAAACCCCACTCTGTCACCCCACT 1914
7823 GATCTTCAAGCATCTTTTACTTTTCACTTCCAGCGTTTCTGGGTGAGCAAAAACAGGAGGAAA 7882
1913 GATCTTCAAGCATCTTTTACTTTTCACTTCCAGCGTTTCTGGGTGAGCAAAAACAGGAGGAAA 1854
7883 ATGCGCGCAAAAACAGGAGGAAATAGGGGACACCGAAATGTTGAAATCTATCTCTCTCTT 7942
1853 ATGCGCGCAAAAACAGGAGGAAATAGGGGACACCGAAATGTTGAAATCTATCTCTCTCTT 1794

QY 7943 TTCAATATTATTGAAGCAATTTATCAGGTTATTGTCTCATGAGCGGATACATATTTGAAT 8002
DB 1793 TTCAATATTATTGAAGCAATTTATCAGGTTATTGTCTCATGAGCGGATACATATTTGAAT 1734
QY 8003 GTATTAGAAAAAATAAACAATAGGGTTCGCGCAATTTCCCGAAAAAGTGCCACCTG 8062
DB 1733 GTATTAGAAAAAATAAACAATAGGGTTCGCGCAATTTCCCGAAAAAGTGCCACCTG 1674
QY 8063 ACGCTAGAAACCAATATTATCATGACATTAACCTATAAAAAATAGCGTATACAGAGGC 8122
DB 1673 ACGCTAGAAACCAATATTATCATGACATTTAACTATATAAATAGCGGTATACAGAGGC 1614
QY 8123 CCTTTGCTTTCAAGAATAATTTCTCATGTTTGAAGCTTTATCATCGATAAGCT 8176
DB 1613 CCTTTGCTTTCAAG-----AATTCTCATGTTTGAAGCTTTATCATCGATAAGCT 1564

RESULT 15

US-08-757-439-1
; Sequence 1, Application US/08757439
; Patent No. 5866371
; GENERAL INFORMATION:
; APPLICANT: BADZIONG, Werner
; APPLICANT: HABERMANN, Paul
; APPLICANT: MOELLER, Joerg
; APPLICANT: ARETZ, Werner
; TITLE OF INVENTION: PROCESS FOR USING THE YEAST ADH II
; TITLE OF INVENTION: PROMOTER SYSTEM FOR THE PRODUCTION OF HETEROLOGOUS
; TITLE OF INVENTION: PROTEINS IN HIGH YIELDS
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/757,439
; FILING DATE: 27-NOV-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 19544233.4
; FILING DATE: 28-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: SANDERCOCK, Colin G.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 18748/303/HOCE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8491 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-757-439-1

Query Match 27.9%; Score 2302; DB 2; Length 8491;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5845 CTGCCTCGCGCTTTTCGGTGATGACGGTGAACACCTCTGACACATGACAGCTCCCGGAGAC 5904
DB 1508 CTGCCTCGCGCTTTTCGGTGATGACGGTGAACACCTCTGACACATGACAGCTCCCGGAGAC 1567

QY 5905 GGTCAAGCTTTGTCTGAAGCGGATCCCGGAGCAGACAAGCCCGTCAAGCGCGCTCAGC 5964
DB 1568 GGTCAAGCTTTGTCTGAAGCGGATCCCGGAGCAGACAAGCCCGTCAAGCGCGCTCAGC 1627
QY 5965 GGGTGTTCGGGGTGTTCGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTGA 6024
DB 1628 GGGTGTTCGGGGTGTTCGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTGA 1687
QY 6025 TACTGCTTTAACTATTCGGGCATCAGAGCAGATTGTACTGAGAGTGACCATATCGGGTGT 6084
DB 1688 TACTGCTTTAACTATTCGGGCATCAGAGCAGATTGTACTGAGAGTGACCATATCGGGTGT 1747
QY 6085 GAAATACCGCAGAGATGCGTAAGAGGAAATACCGGATCAGGGGTCTTCCCGTTCCTCG 6144
DB 1748 GAAATACCGCAGAGATGCGTAAGAGGAAATACCGGATCAGGGGTCTTCCCGTTCCTCG 1807
QY 6145 CTCACCTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 6204
DB 1808 CTCACCTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 1867
QY 6205 GCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGAGGAAAGAAATGAGCAAAA 6264
DB 1868 GCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGAGGAAAGAAATGAGCAAAA 1927
QY 6265 GGCAGCAAAAGCCAGGAAACCGTAAAGAGCGGGTTCGCTGGCGTTTTCATAGGCTC 6324
DB 1928 GGCAGCAAAAGCCAGGAAACCGTAAAGAGCGGGTTCGCTGGCGTTTTCATAGGCTC 1987
QY 6325 CGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCGGACA 6384
DB 1988 CGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCGGACA 2047
QY 6385 GGAATAAAGATACAGCGGTTTCCCTCGTGAAGCTCCCTGCGGCTCTCTCTGTTCCG 6444
DB 2048 GGAATAAAGATACAGCGGTTTCCCTCGTGAAGCTCCCTGCGGCTCTCTCTGTTCCG 2107
QY 6445 ACCCTCGCGTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAAGCGTGGCGCTTCT 6504
DB 2108 ACCCTCGCGTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAAGCGTGGCGCTTCT 2167
QY 6505 CATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTTCGTTTCGCTCCAAAGTGGCGTGT 6564
DB 2168 CATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTTCGTTTCGCTCCAAAGTGGCGTGT 2227
QY 6565 GTGACGAAACCCCGCTTTCAGCGCGCTTCGCGCTTATCCGTAACCTATCGTCTTGAG 6624
DB 2228 GTGACGAAACCCCGCTTTCAGCGCGCTTCGCGCTTATCCGTAACCTATCGTCTTGAG 2287
QY 6625 TCCAAACCCGTTAAGACACGACTTATCCCACTGGCAGCAGCCACTGGTAAACAGGATTAGC 6684
DB 2288 TCCAAACCCGTTAAGACACGACTTATCCCACTGGCAGCAGCCACTGGTAAACAGGATTAGC 2347
QY 6685 AGAGCAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACCTACGCGCTAC 6744
DB 2348 AGAGCAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACCTACGCGCTAC 2407
QY 6745 ACTAGAAGGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAAGA 6804
DB 2408 ACTAGAAGGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAAGA 2467
QY 6805 GTTGGTAGCTTTGATCCGGCAAAACCAACCGCTGTAGCGGTGGTGTGTTTTTTGTTTCG 6864
DB 2468 GTTGGTAGCTTTGATCCGGCAAAACCAACCGCTGTAGCGGTGGTGTGTTTTTTGTTTCG 2527
QY 6865 AGCAGCAGATTACCGCGCAAAAGGATCTCAAGAGATCTCTTTGATCTTTTCTACG 6924
DB 2528 AGCAGCAGATTACCGCGCAAAAGGATCTCAAGAGATCTCTTTGATCTTTTCTACG 2587
QY 6925 GGGTCTGACGCTCAGTGGAAACGAAAACTCAGGTTAAGGGATTTTGTCTAGGATTATCA 6984
DB 2588 GGGTCTGACGCTCAGTGGAAACGAAAACTCAGGTTAAGGGATTTTGTCTAGGATTATCA 2647

Qy		6985	AAAAGGATCTTCACCTAGATCCCTTTTAAATTAAAAATGAAGTTTAAATCAATCTAAAGT	7044
Dd		2648	AAAAGGATCTTCACCTAGATCCCTTTTAAATTAAAAATGAAGTTTAAATCAATCTAAAGT	2707
Qy		7045	ATATAAGTAACAATCTGGTCTGACAGTTAACAATGCTTAATCATAGTAGGAGCACCTACTCA	7104
Dd		2708	ATATAAGTAACAATCTGGTCTGACAGTTAACAATGCTTAATCATAGTAGGAGCACCTACTCA	2767
Qy		7105	GCGATCTGTCTAATTTGGTTCCATCCATAGTTGCCTGACTCCCCTCGTGATAGATAACTACG	7164
Dd		2768	GCGATCTGTCTAATTTGGTTCCATCCATAGTTGCCTGACTCCCCTCGTGATAGATAACTACG	2827
Qy		7165	ATACGGAGGGCTTAACCATCTGCCCCACAGTCTGCAANTGATACCGGAGACCACGCTCA	7224
Dd		2828	ATACGGAGGGCTTAACCATCTGCCCCACAGTCTGCAANTGATACCGGAGACCACGCTCA	2887
Qy		7225	CCGGCTCCAGATTATCAGCAAATAAACACGACGCGGAGAGGCCGACGCGCAGAAGTGCT	7284
Dd		2888	CCGGCTCCAGATTATCAGCAAATAAACACGACGCGGAGAGGCCGACGCGCAGAAGTGCT	2947
Qy		7285	CCTGCACATTTATCCGGCTCCATCCAGTCTAATTTGTTGCCGGGAAGCTAGAGTAAGT	7344
Dd		2948	CCTGCACATTTATCCGGCTCCATCCAGTCTAATTTGTTGCCGGGAAGCTAGAGTAAGT	3007
Qy		7345	AGTTCGGCAGTTAATAGTTTTGGCAACAGTTGTTGCCATTCCTCGAGCATCGTGTTGTC	7404
Dd		3008	AGTTCGGCAGTTAATAGTTTTGGCAACAGTTGTTGCCATTCCTCGAGCATCGTGTTGTC	3067
Qy		7405	CGCTCGTCGTTTGATATGGCTTCATTCAAGCTCCGGTTCCAAACGATCAAGCGCAGATTACA	7464
Dd		3068	CGCTCGTCGTTTGATATGGCTTCATTCAAGCTCCGGTTCCAAACGATCAAGCGCAGATTACA	3127
Qy		7465	TGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCTTCGGTCTCCGATCGTTGTCAGA	7524
Dd		3128	TGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCTTCGGTCTCCGATCGTTGTCAGA	3187
Qy		7525	AGTAAGTTGCGCGCAGTGTTATCACTCATGGTTATGACACACTGCAATAATCTCTTACT	7584
Dd		3188	AGTAAGTTGCGCGCAGTGTTATCACTCATGGTTATGACACACTGCAATAATCTCTTACT	3247
Qy		7585	GTCAATGCATCCGTAAGATGCTTTTGTGACTGGTGAGTACTCAACCAAGTCAATCTGA	7644
Dd		3248	GTCAATGCATCCGTAAGATGCTTTTGTGACTGGTGAGTACTCAACCAAGTCAATCTGA	3307
Qy		7645	GAATAGTGTATCGGCGACGAGTTGCTCTTCCCGCGCTCAACACGGGATAATACCGCG	7704
Dd		3308	GAATAGTGTATCGGCGACGAGTTGCTCTTCCCGCGCTCAACACGGGATAATACCGCG	3367
Qy		7705	CCACATAGCAGAACTTTTAAAGTGCTCATCATTTGGAAAAAGCTTCTCGGGCGGAAAACTC	7764
Dd		3368	CCACATAGCAGAACTTTTAAAGTGCTCATCATTTGGAAAAAGCTTCTCGGGCGGAAAACTC	3427
Qy		7765	TCAAGGATCTTAACGCTGTTGAGATCAGTTCGATGTAAACCACTCGTCACCCCACTGA	7824
Dd		3428	TCAAGGATCTTAACGCTGTTGAGATCAGTTCGATGTAAACCACTCGTCACCCCACTGA	3487
Qy		7825	TCCTTCAGCATCTTTTACTTTTCCACGAGGTTTCTGGGTGAGCAAAAAACAGGAAGCAAAAT	7884
Dd		3488	TCCTTCAGCATCTTTTACTTTTCCACGAGGTTTCTGGGTGAGCAAAAAACAGGAAGCAAAAT	3547
Qy		7885	GCCGCAAAAAAGGGAATAAGGCGACACGGAAATGTTGTAATCTCATACTCTTCCTTTT	7944
Dd		3548	GCCGCAAAAAAGGGAATAAGGCGCGACACGGAAATGTTGTAATCTCATACTCTTCCTTTT	3607
Qy		7945	CAATATTATTGAAGCATTTATCAGGTTATTGCTCATGAGCGGATACATATTGTAATGT	8004
Dd		3608	CAATATTATTGAAGCATTTATCAGGTTATTGCTCATGAGCGGATACATATTGTAATGT	3667
Qy		8005	ATTTAGAAAAATAAACAATAGGGGTTCCGCGCATTTTCCCGAAAAAGTGCCACTTGAC	8064
Dd		3668	ATTTAGAAAAATAAACAATAGGGGTTCCGCGCATTTTCCCGAAAAAGTGCCACTTGAC	3727
Qy		8065	CTCTABACBACCATATTATTATCATCATTAATTAACCTTATATAAAATACGGCTATACAGAGGCC	8124

Db	3728	GTCTAAGAACCATTTATCATGACATTAAACCTATAAAATAGGGGTATCAGAGGCC	3787
Qy	8125	TTTCGTCCTTCAAGAAATTAATTC	8146
Db	3788	TTTCGTCCTTCAAGAAATTAATTC	3809

Search completed: September 24, 2004, 10:08:54
Job time : 389.227 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2004, 17:04:19 ; Search time 12503.4 seconds
(without alignments)
19682.126 Million cell updates/sec

Title: US-10-030-390-3
Perfect score: 8241
Sequence: 1 agatctaacatccaagaagc.....aaataacagttattatcg 8241

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_estc:*
- 9: gb_estl:*
- 10: gb_est2:*
- 11: gb_est3:*
- 12: gb_est4:*
- 13: gb_est5:*
- 14: gb_est6:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rod:*
- 26: em_gss_phg:*
- 27: em_gss_vrl:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1055.8	12.8	1070	9 AJ281552	AJ281552 4A3A-P6F1
2	993.2	11.9	1013	12 BM438846	BM438846 IplYr0015
3	924	11.2	1004	9 AJ281480	AJ281480 4A3A-P4G8
4	915.8	11.1	973	14 CD458281	CD458281 Fg08_08hC

C 5	885.2	10.7	917	14	CD458286	CD458286 Fg08_09a0
C 6	872.6	10.6	935	12	BG838279	BG838279 Gc01_10e0
C 7	865	10.5	918	14	CD459092	CD459092 Fg08_08e0
C 8	837.8	10.2	841	9	AL042026	AL042026 DKFZp434E
C 9	829.6	10.1	872	14	CD459085	CD459085 Fg08_08d0
C 10	821.4	10.0	854	12	BM438950	BM438950 IplYr0049
C 11	819.2	9.9	870	14	CD458333	CD458333 Fg08_09e0
C 12	811.2	9.8	1126	28	BZ577702	BZ577702 msh2_5533
C 13	804.8	9.8	819	14	CD649375	CD649375 CvgHd0008
C 14	794.4	9.6	1011	28	BZ576726	BZ576726 msh2_5071
C 15	785.4	9.5	1574	28	BZ572566	BZ572566 msh2_2693
C 16	779.4	9.5	1073	14	CF269652	CF269652 Pcy1C01d8
C 17	777.8	9.4	800	9	AJ281449	AJ281449 4A3A-F4D5
C 18	771.6	9.4	1336	28	BZ575810	BZ575810 msh2_4637
C 19	771.2	9.4	1370	28	BZ571721	BZ571721 msh2_2025
C 20	770.4	9.4	789	14	CD280920	CD280920 G44224_42
C 21	770.4	9.3	1249	28	BZ572284	BZ572284 msh2_2572
C 22	763.8	9.3	966	28	BZ570738	BZ570738 msh2_1513
C 23	760	9.2	780	13	BQ825693	BQ825693 1030I29B0
C 24	753.8	9.1	759	14	CD279661	CD279661 G43818_35
C 25	752.2	9.1	1463	28	BZ571475	BZ571475 msh2_1906
C 26	745.4	9.0	844	28	BZ574513	BZ574513 msh2_3706
C 27	745	9.0	786	14	CD458721	CD458721 Fg08_04e0
C 28	741.8	9.0	1003	28	BZ576686	BZ576686 msh2_5053
C 29	740.6	9.0	966	28	BZ575002	BZ575002 msh2_4255
C 30	740.4	9.0	832	12	BQ923768	BQ923768 602825893
C 31	739.2	9.0	914	28	BZ569398	BZ569398 pac82-164
C 32	737.8	9.0	741	14	CD279174	CD279174 G44221_83
C 33	733.2	8.9	811	29	ATH517156	AJ517156 Arabidops
C 34	730.6	8.9	950	28	BZ571129	BZ571129 msh2_1741
C 35	729.2	8.8	998	28	BZ576702	BZ576702 msh2_5060
C 36	728	8.8	954	9	AL044364	AL044364 DKFZp434C
C 37	726.8	8.8	730	14	CD281097	CD281097 G44224_38
C 38	726.2	8.8	1067	9	AU081137	AU081137 AU081137
C 39	726	8.8	840	28	BZ571995	BZ571995 msh2_220.
C 40	725.8	8.8	729	14	CD279546	CD279546 G44222_81
C 41	725.4	8.8	842	12	B1687610	B1687610 603312586
C 42	724.8	8.8	728	14	CD279322	CD279322 G44223_13
C 43	723	8.8	881	28	BZ577222	BZ577222 msh2_5308
C 44	722.8	8.8	726	14	CD281811	CD281811 G4471_30
C 45	722.8	8.8	1394	28	BZ580042	BZ580042 msh2_908.

ALIGNMENTS

RESULT 1	AJ281552	1070 bp	mRNA	linear	EST 30-JUN-2000
AJ281552	4A3A-P6F11-F	Anopheles gambiae	immune competent	4A3A	Anopheles
LOCUS	gambiae cDNA clone 4A3A-P6F11	mRNA	sequence.		
DEFINITION	AJ281552	gambiae cDNA clone 4A3A-P6F11	mRNA	sequence.	
ACCESSION	AJ281552	GI:6929432			
VERSION	AJ281552.1	GI:6929432			
KEYWORDS	EST.				
SOURCE	Anopheles gambiae	(African malaria mosquito)			
ORGANISM	Anopheles gambiae				
REFERENCE	1	(bases 1 to 1070)			
AUTHORS	Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C., Dorchug, M., Schultz, J., Benes, V., Bork, P., Ansoorge, W., Soares, M.B. and Kafatos, F.C.				
TITLE	Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	97 (12), 6619-6624	(2000)		
MEDLINE	20300950				
PUBMED	10841561				
COMMENT	Contact: Dimopoulos G Fotis C. Kafatos laboratory European Molecular Biology Laboratory Meyerhofstrasse 1, 69117 Heidelberg, Germany.				

FEATURES
source

Location/Qualifiers
1. .1070
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="4a r/r"
/db_xref="taxon:7165"
/clone="4A3A-P6F11"
/cell_line="immune competent 4A3A"
/lab_host="E. coli DH10B"
/clone_vector="Anopheles gambiae immune competent 4A3A"
/note="Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site 1: EcoRI; Site 2: NotI; sequenced from forward priming site which reads from the 3' end of the cDNA. The 4A3A is a directionally cloned and normalized cDNA library that was constructed from the 4A3A cell line oligo-T primed cDNA according to: Bonaldo, Lennon & Soares (1996): Normalization and Subtraction: Two approaches to Facilitate Gene Discovery, Genome Research 6, 791-806."

ORIGIN

Query Match 12.8%; Score 1055.8; DB 9; Length 1070;
Best Local Similarity 99.7%; Pred. No. 5e-224;
Matches 1068; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 6301 TTGCTGGGCTTTTCATAGCTCGGCGCCCTGACGAGCATCACAAAATCGACGCTCA 6360
DB 1 TTGCTGGGCTTTTCATAGCTCGGCGCCCTGACGAGCATCACAAAATCGACGCTCA 60
QY 6361 AGTCAGAGGTGGGAAACCCGACAGGACTATAAGATACACAGGCTTTCCCTCGAAGC 6420
DB 61 AGTCAGAGGTGGGAAACCCGACAGGACTATAAGATACACAGGCTTTCCCTCGAAGC 120
QY 6421 TCCCTCGTGCCTCTCTGTTCCGACCTGCGGCTTACCGGATACCTGCGGCTTTCTC 6480
DB 121 TCCCTCGTGCCTCTCTGTTCCGACCTGCGGCTTACCGGATACCTGCGGCTTTCTC 180
QY 6481 CTTTCGGGAAGCTGGGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTTGGGTAG 6540
DB 181 CTTTCGGGAAGCTGGGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTTGGGTAG 240
QY 6541 GTGCTTGCCTCAAGCTGGGCTGTGTGACGACCCCGCTTACCGGCTTACCGGCTTGC 6600
DB 241 GTGCTTGCCTCAAGCTGGGCTGTGTGACGACCCCGCTTACCGGCTTACCGGCTTGC 300
QY 6601 TTATCCGGTAACCTATCGTCTTGTAGTCCAAACCCGCTAAGACACACACTTATCGC 6660
DB 301 TTATCCGGTAACCTATCGTCTTGTAGTCCAAACCCGCTAAGACACACACTTATCGC 360
QY 6661 CGAGCACTGTACAGGATAGCAGAGGATGTAGGCGGCTCTACAGTTCTTG 6720
DB 361 CGAGCACTGTACAGGATAGCAGAGGATGTAGGCGGCTCTACAGTTCTTG 420
QY 6721 AAGTGGTGGCTAACTACGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTG 6780
DB 421 AAGTGGTGGCTAACTACGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTG 480
QY 6781 AAGCAGTTACTTCGGAAGAGGTGTAGTCTTGTATCGGCGAACAACCCAGCT 6840
DB 481 AAGCAGTTACTTCGGAAGAGGTGTAGTCTTGTATCGGCGAACAACCCAGCT 540
QY 6841 GGTAGCGTGGT 6900
DB 541 GGTAGCGTGGT 600
QY 6901 GAAGATCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGACGAAATCTCAGTTAA 6960
DB 601 GAAGATCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGACGAAATCTCAGTTAA 660
QY 6961 GGGATTTTGGTCATGAGATTACAAAGAGGATCTTCACCTAGATCTTTTAAATTAATAA 7020
DB 661 GGGATTTTGGTCATGAGATTACAAAGAGGATCTTCACCTAGATCTTTTAAATTAATAA 720
QY 7021 TGAAGTTTAAATCAATCTAAAGTATATATAGTAAACTTTGGTCTGACAGTTTACCAATGC 7080

DB 721 TGAAGTTTAAATCAATCTAAAGTATATATAGTAAACTTTGGTCTGACAGTTTACCAATGC 780
QY 7081 TTATCAGTGAAGCACCCTATCTCAGCGATCTGTCTATTTTCGTTTCTATCCATAGTTCCCTGA 7140
DB 781 TTATCAGTGAAGCACCCTATCTCAGCGATCTGTCTATTTTCGTTTCTATCCATAGTTCCCTGA 840
QY 7141 CTCCCTCTCTGTAGTAACATACGATACGGAGGGCTTACCATCTGCGCCAGTCTGCA 7200
DB 841 CTCCCTCTCTGTAGTAACATACGATACGGAGGGCTTACCATCTGCGCCAGTCTGCA 900
QY 7201 ATGATACCGGAGAGCACCAGCTCACCAGGCTCCAGATTTATCAGCAATAAACACGACGCC 7260
DB 901 ATGATACCGGAGAGCACCAGCTCACCAGGCTCCAGATTTATCAGCAATAAACACGACGCC 960
QY 7261 GGAAGGCCCGAGCGGAGAGTGTCTGCACTTTATCGGCTCCATCCAGTCTTAAAT 7320
DB 961 GGAAGGCCCGAGCGGAGAGTGTCTGCACTTTATCGGCTCCATCCAGTCTTAAAT 1019
QY 7321 TGTTCGGGGAAGCTAGAGTAGTGTTCGCCAGTTAATAGTTTGGCAAC 7371
DB 1020 TGTTCGGGGAAGCTAGAGTAGTGTTCGCCAGTTAATAGTTTGGCAAC 1070

RESULT 2
BM438846/c

LOCUS Iplv00157 Liver cDNA library Ictalurus punctatus cDNA 5', mRNA
DEFINITION BM438846 1013 bp mRNA linear EST 31-JAN-2002
ACCESSION BM438846
VERSION BM438846.1 GI:18460568
KEYWORDS EST.
SOURCE Ictalurus punctatus (channel catfish)
ORGANISM Ictalurus punctatus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
Ictaluridae; Ictalurus.
1 (bases 1 to 1013)
Feng, J., Kucuktas, H., Kocabas, A., Li, P. and Liu, Z.
Transcriptome of channel catfish (Ictalurus punctatus): initial
analysis of expressed sequence tags from the liver
Unpublished (2002)
Contact: Liu ZJ
JOURNAL The Fish Molecular Genetics and Biotechnology Laboratory,
Department of Fisheries and Allied Aquacultures and Program of Cell
and Molecular Biosciences
Auburn University, Auburn, AL 36849, USA
203 Swingle Hall, Auburn University, Auburn, AL 36849, USA
Tel: 334 844 4054
Fax: 334 844 9208
Email: zliu@acesag.auburn.edu
Seq primer: M13 Reverse.
Location/Qualifiers
1. .1013
/organism="Ictalurus punctatus"
/mol_type="mRNA"
/db_xref="taxon:7998"
/clone_lib="Liver cDNA library"
/note="Organ: Liver; Vector: pSport1; Site_1: NotI;
Site_2: SalI"

FEATURES
source

1. .1013
/organism="Ictalurus punctatus"
/mol_type="mRNA"
/db_xref="taxon:7998"
/clone_lib="Liver cDNA library"
/note="Organ: Liver; Vector: pSport1; Site_1: NotI;
Site_2: SalI"

ORIGIN

Query Match 11.9%; Score 983.2; DB 12; Length 1013;
Best Local Similarity 99.3%; Pred. No. 7.5e-208;
Matches 1007; Conservative 1; Mismatches 4; Indels 2; Gaps 2;
QY 6874 ATTACCGCAGAAAAAGGATCTCAAGAGATCTTTGATCTTTTCTACGGGTCTGAC 6933
DB 1013 ATTACCGCAGAAAAAGGATCTCAAGAGATCTTTGATCTTTTCTACGGGTCTGAC 954
QY 6934 GCTCAGTGAAGCAAACTCAGTTAAGGATTTTGGTCATGAGATTATCAAAAGGATC 6993
DB 953 GCTCAGTGAAGCAAACTCAGTTAAGGATTTTGGTCATGAGATTATCAAAAGGATC 894


```

Db 481 GAACAGATATTGGTATCTGCGCTCTGCTGAAGCAGTTACCTTCGGAAAAAGAGTTGGTA 540
Qy 6812 GCTCTTGATCGGCAACCAACACCGCTGGTAGCGGTGTTTGTGTCAGCAGC 6871
Db 541 GCTCTTGATCGGCAACCAACACCGCTGGTAGCGGTGTTTGTGTCAGCAGC 600
Qy 6872 AGATTACCGCAGAAAAAGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCG 6931
Db 601 AGATTACCGCAGAAAAAGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCG 660
Qy 6932 AGCTCAGTGAACGAAACTCAGTTAAGGATTTTGTTCATGAGATATCAAAAGCA 6991
Db 661 AGCTCAGTGAACGAAACTCAGTTAAGGATTTTGTTCATGAGATATCAAAAGCA 720
Qy 6992 TCTTCACCTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATG 7051
Db 721 TCTTCACCTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATG 780
Qy 7052 AGTAACTGCTGACAGTACCAATGCTTATCAGTGAAGGACCTATCTCAGCGATCT 7111
Db 781 AGTAACTGCTGACAGTACCAATGCTTATCAGTGAAGGACCTATCTCAGCGATCT 839
Qy 7112 GTCTATTTGCTTATCCTAGTGTGCTGACTCCCGCTCGTGTAGATAACTAGTACGGG 7171
Db 840 GTCTATTTGCTTATCCTAGTGTGCTGACTCCCGCTCGTGTAGATAACTAGTACGGG 898
Qy 7172 AGGCTTACCTGCGCCCGCTGCTCAATGATACCGGAGACCACTCAGCGGCTC 7231
Db 899 AGGCTTACCTGCGCCCGCTGCTCAATGATACCGGAGACCACTCAGCGGCTC 955
Qy 7232 CAGATTTATCAGCAATAAACCAGCAGCCGGAAGGCGGCGGCGAGCAAGTG 7282
Db 956 CAGATTT-TCAGCAATAAACCAGCAGCCGGAAGGCGGCGGCGAGCAAGTG 1004

```

RESULT 4

```

CD458281/c 973 bp mRNA linear EST 03-JUN-2003
LOCUS Fg08_08h09_R Fg08_AAFc ECORC Fusarium graminearum complex_substrate
DEFINITION Gibberella zeae cDNA clone Fg08_08h09, mRNA sequence.
ACCESSION CD458281
VERSION CD458281.1 GI:31373021
KEYWORDS EST.
SOURCE Gibberella zeae
ORGANISM Gibberella zeae
REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
AUTHORS Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
1 (bases 1 to 973)
Watson, R.J., Heys, R., Chapados, J., Couroux, P., Harris, L.J.,
Hattori, J., Lacroix, C., Ouellet, T., Robert, L.S., Singh, J.A.,
Spotts, D. and Finker, N.A.
A cDNA library prepared from Fusarium graminearum grown on a
complex plant substrate
Contact: Watson, Robert.J.
Unpublished (2003)
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-Food Canada
Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
CANADA
Tel: (613) 759-1655
Fax: (613) 759-1701
Email: watsonrj@agr.gc.ca.
Location/Qualifiers
1. 973
/organism="Gibberella zeae"
/mol_type="mRNA"
/strain="DAOM 180378"
/db_xref="taxon:5518"
/clone="Fg08_08h09"
/tissue_type="Mycelium"
/dev_stage="Asexual"
/lab_host="E. coli DH10B"

```

FEATURES

source

```

/clone_lib="Fg08_AAFc ECORC Fusarium graminearum complex_s
ubstrate"
/notes="Vector: pBluescript II+; Site 1: EcoRI; Site 2:
XhoI; Fusarium graminearum grown on a complex plant
substrate-- wheat leaves treated to remove most of the low
molecular weight, water-soluble components."
ORIGIN
Query Match 11.1%; Score 915.8; DB 14; Length 973;
Best Local Similarity 98.1%; Pred. No. 7.8e-193;
Matches 931; Conservative 5; Mismatches 12; Indels 1; Gaps 1;
Qy 7123 TCATTCATAGTTCCTGACCTCCCGTCTGTGATAGTAACATACGATACGGAGGAGCTTACCA 7182
Db 965 TCATTCATAGTTCCTGACCTCCCGTCTGTGATAGTAACATACGATACGGAGG-RRGCTTACCA 907
Qy 7183 TCTGGCCCGGAGTCTGCAATGATACCGGAGACCAACGCTCACCGCTCCAGATTTATCA 7242
Db 906 TCTGGCCCGGAGTCTGCAATGATACCGGAGACCAACGCTCACCGCTCCAGATTTATCA 847
Qy 7243 GCATTAACACGACCCCGGAGGAGGCGGAGGAGGAGTGTCTGCACTTTATCCGCC 7302
Db 846 GCATTAACACGACCCCGGAGGAGGCGGAGGAGTGTCTGCACTTTATCCGCC 787
Qy 7303 TCCATCCAGTCTATTAATTTGTTGCCGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGT 7362
Db 786 TCCATCCAGTCTATTAATTTGTTGCCGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGT 727
Qy 7363 TTGGGCAAGGTTGTCCTGATTCGAGCATGCTGTGTGTCACGCTCGCTGTTGGTATG 7422
Db 726 TTGGGCAAGGTTGTCCTGATTCGATCAGGCACTGCTGTGTGTCACGCTCGCTGTTGGTATG 667
Qy 7423 GCTTCATTCAGTCTCCGTTCCCAACGATCAAGCGGAGTTACATATCCCCCATGTTGTGC 7482
Db 666 GCTTCATTCAGTCTCCGTTCCCAACGATCAAGCGGAGTTACATATCCCCCATGTTGTGC 607
Qy 7483 AAAAAAGCGTTAGTCTTCGCTTCGATCGTGTGTGTCAGAGTAAGTTGGCGCGAGTG 7542
Db 606 AAAAAAGCGTTAGTCTTCGCTTCGATCGTGTGTGTCAGAGTAAGTTGGCGCGAGTG 547
Qy 7543 TTATCACTCATGTTATGTCAGCACTGTCATTAATTTCTTACTGTGTCATGCCATCCGTAAGA 7602
Db 546 TTATCACTCATGTTATGTCAGCACTGTCATTAATTTCTTACTGTGTCATGCCATCCGTAAGA 487
Qy 7603 TGCTTTTCTGAGTGTGAGTACTCAACCAAGTCAATCTGAGAAATAGTATGCGGCA 7662
Db 486 TGCTTTTCTGAGTGTGAGTACTCAACCAAGTCAATCTGAGAAATAGTATGCGGCA 427
Qy 7663 CCAGTTGCTCTTGCCCGCGTCAACAGGGGATAATACCGGCCACATAGCAGAACTTTA 7722
Db 426 CCAGTTGCTCTTGCCCGCGTCAATCGGGATATACCGGCCACATAGCAGAACTTTA 367
Qy 7723 AAGTCTCATCATTTGGAAAAAGCTTCTTCGGGGGAAAAAAGTCTCAGGATCTTACCGCTG 7782
Db 366 AAGTCTCATCATTTGGAAAAAGCTTCTTCGGGGGAAAAAAGTCTCAGGATCTTACCGCTG 307
Qy 7783 TTGAGATCCAGTTCGATGTAACCCACTCGTGCAACCCACTGATCTTCAGCATCTTTTACT 7842
Db 306 TTGAGATCCAGTTCGATGTAACCCACTCGTGCAACCCACTGATCTTCAGCATCTTTTACT 247
Qy 7843 TTCACGAGGTTCTCGGTGAGCAAAAAAGGAGGAAAAATGCGCAAAAAAGGAAATA 7902
Db 246 TTCACGAGGTTCTCGGTGAGCAAAAAAGGAGGAAAAATGCGCAAAAAAGGAAATA 187
Qy 7903 AGGGGACACGGAATTTGTAATCTCATCTCTTCCCTTTTCAATATTTTGAAGCAT 7962
Db 186 AGGGGACACGGAATTTGTAATCTCATCTCTTCCCTTTTCAATATTTTGAAGCAT 127
Qy 7963 TATCAGGTTTATGTTCTCATGAGCGATACATATTTGATGATTTAGAAAAATAACAA 8022
Db 126 TATCAGGTTTATGTTCTCATGAGCGATACATATTTGATGATTTAGAAAAATAACAA 67
Qy 8023 ATAGGGTTCCCGGACACATTTTCCCGAAAAAGTGCCACCTGAGCTTAAG 8071

```

QY	7474	ATGTTGTGCAAAAAGCGGTAGCTCCTTCGGTCCCTCCGACTCGTTGCAGAAGTAAGTTG	7533
Db	617	ATGTTGTGCAAAAAGCGGTAGCTCCTTCGGTCCCTCCGACTCGTTGCAGAAGTAAGTTG	558
QY	7534	GCCGACGTATTACACTCATGTTATGGCAGCACTGCATAAATTTCTTACTGTTCATGCCA	7593
Db	557	GCCGACGTATTACACTCATGTTATGGCAGCACTGCATAAATTTCTTACTGTTCATGCCA	498
QY	7594	TCCGTAAAGATGCTTTTTCTGTGACTGGTGAGTACTCAACCAAAGTCATTCTGAGGAATAGTGT	7653
Db	497	TCCGTAAAGATGCTTTTTCTGTGACTGGTGAGTACTCAACCAAAGTCATTCTGAGGAATAGTGT	438
QY	7654	ATCGCGGACCGAGTGTCTCTTCGCCGGCGTCAACAGGGATAAATACCGCGCACCATAGC	7713
Db	437	ATCGCGGACCGAGTGTCTCTTCGCCGGCGTCAAATACGGGATAAATACCGCGCACCATAGC	378
QY	7714	AGAACTTTAAAAAGTGTCTCATCTATGGAAAAACGTTCTTCGGGGCGAAAAAATCTCAAGGATC	7773
Db	377	AGAACTTTAAAAAGTGTCTCATCTATGGAAAAACGTTCTTCGGGGCGAAAAAATCTCAAGGATC	318
QY	7774	TTACCGCTGTTGAGATCCAGTTCGATGTAAACCACATCGTGCACCAACTGATCTTCAGCA	7833
Db	317	TTACCGCTGTTGAGATCCAGTTCGATGTAAACCACATCGTGCACCAACTGATCTTCAGCA	258
QY	7834	TCTTTTACTTTTACCAGCGTTTCTCGGTGAGCAAAAACAGGAAGGCAAAATGCGCAAAA	7893
Db	257	TCTTTTACTTTTACCAGCGTTTCTCGGTGAGCAAAAACAGGAAGGCAAAATGCGCGAAA	198
QY	7894	AAGGGAATAAGCGGCACACGAAATGTTGAATCTCATCTACTCTCTTTTCAATATTAT	7953
Db	197	AAGGGAATAAGCGGCACACGAAATGTTGAATCTCATCTACTCTCTTTTCAATATTAT	138
QY	7954	TGAAGCATTTATCAGGGTTATTCTCTCATGAGCGGATACATATTTCGAATGTTTTAGAAA	8013
Db	137	TGAAGCATTTATCAGGGTTATTCTCTCATGAGCGGATACATATTTCGAATGTTTTAGAAA	78
QY	8014	AATAAACAAATAGGGTTCCGCGCATTTCCCGCAAAATGTCACCTGACGCTCTAAG	8071
Db	77	AATAAACAAATAGGGTTCCGCGCATTTCCCGCAAAATGTCACCTAATTTGTTAAG	20
RESULT	6		
EG838279/c			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			

Db 148 ATATTGAAGCAATTATACAGGGTTATATGTCTCATGAGCGGATACATATATTTGAATGTTATT 89
 QY 8009 AGAAAAATAAAACAATAGAGGGTTCGGCGCACATTTCCCGAAAAGTGCACCTGACGTCT 8068
 Db 88 AGAAAAATAAACAATAGAGGGTTCGGCGCACATTTCCCGAAAAGTGCACCTGACCTAAATGTT 29
 QY 8069 AAG 8071
 Db 28 AAG 26
 RESULT 7
 CD459092/c
 LOCUS
 DEFINITION
 FG08_08e02_R FG08_AAPC_ECORC Fusarium graminearum complex_substrate
 Gibberella zeae cDNA clone FG08_08e02, mRNA sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 CD459092.1 GI:31373832
 EST.
 Gibberella zeae
 Gibberella zeae
 Gibberella zeae
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Hypocreomycetidae; Hypocreales, Nectriaceae; Gibberella.
 1 (bases 1 to 918)
 Watson,N.J., Heyes,R., Chapados,J., Couroux,P., Harris,L.J.,
 Hattori,J., Macroix,C., Ouellet,T., Robert,D.S., Singh,J.A.,
 Spott,D. and Tinker,N.A.
 A cDNA library prepared from Fusarium graminearum grown on a
 complex plant substrate
 Unpublished (2003)
 Contact: Watson, Robert.J.
 Eastern Cereal and Oilseed Research Centre
 Agriculture and Agri-food Canada
 Bldg. 20, Central Experimental Farm, Ottawa, Ontario, KIA 0C6,
 CANADA
 Tel: (613) 759-1655
 Fax: (613) 759-1701
 Email: watsonrj@agr.gc.ca.
 Location/Qualifiers
 1..918
 /organism="Gibberella zeae"
 /mol_type="mRNA"
 /strain="DAOM 180378"
 /db_xref="taxon:5518"
 /clone="FG08_08e02"
 /tissue_type="Mycelium"
 /dev_stage="Asexual"
 /lab_host="E. coli DH103"
 /clone_lib="FG08_AAPC_ECORC_Fusarium_graminearum_complex_s
 ubstrate"
 /note="Vector: pBluescript II+; Site_1: EcoRI; Site_2:
 XhoI; Fusarium graminearum grown on a complex plant
 substrate-- wheat leaves treated to remove most of the low
 molecular weight, water-soluble components."
 ORIGIN
 Query Match 10.5%; Score 865; DB 14; Length 918;
 Best Local Similarity 98.8%; Pred. No. 1,6e-181;
 Matches 890; Conservative 2; Mismatches 7; Indels 2; Gaps 2;
 QY 7173 GGGCTTACATCTGCCCGCAGTGTGCAATGA-TACCGCGAGACCCAGCTCACCGGCTC 7231
 Db 918 GGGCTTACATCTGCCCGCAGTGTGCAATGATTACCGCGAGMCCAGCTCACCGGCTC 859
 QY 7232 CAGATTATTACGCAATAAACAGCCAGCCGGAAGGGCCGACGAGAGTGTCTCTGCAA 7291
 Db 858 CAGATTATTACGCAATAAACAGCCAGCCGGAAGGGCCGACGAGAGTGTCTCTGCAA 799
 QY 7292 CTTTATCGGCTCCATCCAGTCTATTAAATGTTCCGGGAAGCTAGTAGTAGTTCGC 7351
 Db 798 CTTTATCGGCTCCATCCAGTCTATTAAATGTTCCGGGAAGCTAGTAGTAGTTCGC 739
 QY 7352 CAGTTTAAATAGTTTGGCGAACCGTTGTGTCATGTGTCAGGATCGTGGTGTGTCAGCTCGT 7411

738 CAGTTAATAGTTTGGCGCAAGCTTGTGGCAATGCTACAGGCATCGTGGTGTACGCTCGT 679
7412 CGTTTGGTATGGCTTCAATTGAGTCCCGTCCCAACGATCAAGCGAGTTACATGATCCC 7471
678 CGTTTGGTATGGCTTCAATTGAGTCCCGTCCCAACGATCAAGCGAGTTACATGATCCC 619
7472 CCATGTTGTGCAAAAAGCGGTAGCTCCTTCGGTCTCCGATCGTGTGTCAGAAAGTAAGT 7531
618 CCATGTTGTGCAAAAAGCGGTAGCTCCTTCGGTCTCCGATCGTGTGTCAGAAAGTAAGT 559
7532 TGGCCGAGTGTATCACTCA-TGGTTATGGCAGCACTGCATATATCTCTTACTGTCAATG 7590
558 TGGCCGAGTGTATCACTCACTGAGTGTATGGCAGCACTGCATATATCTCTTACTGTCAATG 499
7591 CCATCCGTAAAGATGCTTTTCTGTGACCTGGTGAGTACTCAACCAAGTCACTCTCAGATAG 7650
498 CCATCCGTAAAGATGCTTTTCTGTGACCTGGTGAGTACTCAACCAAGTCACTCTCAGATAG 439
7651 TGTATGCGGCGACCGAGTGTCTCTTTCGCCGCGCTCAACACGGGATATACCGGCGCACAT 7710
438 TGTATGCGGCGACCGAGTGTCTCTTTCGCCGCGCTCAACACGGGATATACCGGCGCACAT 379
7711 AGCAGAACTTTAAAGTGTCTCATCTGGAAGAGTCTTTCGGGGGCAAAACTCTCAAGG 7770
378 AGCAGAACTTTAAAGTGTCTCATCTGGAAGAGTCTTTCGGGGGCAAAACTCTCAAGG 319
7771 ATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGCACCCAACTGATCTTCA 7830
318 ATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGCACCCAACTGATCTTCA 259
7831 GCATCTTTTACTTTCACAGCGTTCCTGGTGAGCAAGAAACAGAGAGGCAAAATGCCGA 7890
258 GCATCTTTTACTTTCACAGCGTTCCTGGTGAGCAAGAAACAGAGAGGCAAAATGCCGA 199
7891 AAAAAGGCAATAAGGGCGACACGAAATGTTGAATGTAATCTACTCTCTCTTTTCAATAT 7950
198 AAAAAGGCAATAAGGGCGACACGAAATGTTGAATGTAATCTACTCTCTCTTTTCAATAT 139
7951 TATTGAAGCAATTATCAGGTTATGTTCTCATGCGGATACATATTTGAATGTAATTAG 8010
138 TATTGAAGCAATTATCAGGTTATGTTCTCATGCGGATACATATTTGAATGTAATTAG 79
8011 AAAAATAACAAATAGGGTTCGCGCACATTTTCCGGAAGTGCACCTGACGCTAA 8070
78 AAAAATAACAAATAGGGTTCGCGCACATTTTCCGGAAGTGCACCTGACCTAAATGTA 19
8071 G 8071
18 G 18

RESULT 8
AL042026/c
LOCUS
DEFINITION DKF2p434E11.1 r1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION AL042026
VERSION AL042026.1 GI:5421372
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 841)
Poustka, A., Klein, M., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
EST (Poustka, et al.)
Unpublished (1999)
Contact: MIPS
MIPS
Ingolstaedter Landstr. 1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No si sequence available.
This clone (DKFZp434E11.1) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1..841
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp434E11.1"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN
Query Match 10.2%; Score 837.8; DB 9; Length 841;
Best Local Similarity 99.8%; Pred No. 1.8e-175;
Matches 339; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 7185 TGGCCCCAGTGTGCAATGATACCGGAGACCCAGCTCACCGGCTCCAGATTTATCAGC 7244
DB 841 TGGCCCCAGTGTGCAATGATACCGGAGACCCAGCTCACCGGCTCCAGATTTATCAGC 782
QY 7245 AATAAACAGCAGCGGAGGCGGAGGCGGAGGAGTGTCTGCAACTTTATCCGCTC 7304
DB 781 AATAAACAGCAGCGGAGGCGGAGGCGGAGGAGTGTCTGCAACTTTATCCGCTC 722
QY 7305 CATCCAGTCTATTAAATTGTTCCGCGGAAGCTAGAGTAGTTCGCCAGTAAATAGTTT 7364
DB 721 CATCCAGTCTATTAAATTGTTCCGCGGAAGCTAGAGTAGTTCGCCAGTAAATAGTTT 662
QY 7365 GCGCAACGTTGTGTCATTTGCTGAGGATCGTGTGTGTCAGCTCGTGTGGTATGCG 7424
DB 661 GCGCAACGTTGTGTCATTTGCTGAGGATCGTGTGTGTCAGCTCGTGTGGTATGCG 602
QY 7425 TTCAATTCAGCTCCGCTTCCCAACGATCAAGCGAGTTACATGATCCCCCATGTTGTGCA 7484
DB 601 TTCAATTCAGCTCCGCTTCCCAACGATCAAGCGAGTTACATGATCCCCCATGTTGTGCA 542
QY 7485 AAAAGCGTTAGCTTCCTTCGCTTCGATCGTGTGTCAGAAAGTAAGTTCGCCAGTGT 7544
DB 541 AAAAGCGTTAGCTTCCTTCGCTTCGATCGTGTGTCAGAAAGTAAGTTCGCCAGTGT 482
QY 7545 ATCACTCATGTTATGGCAGCACTGCATTAATTCCTTACTCTCATGCTCCATCGTAAGATG 7604
DB 481 ATCACTCATGTTATGGCAGCACTGCATTAATTCCTTACTCTCATGCTCCATCGTAAGATG 422
QY 7605 CTTTCTCTGACTGGTGTAGTACTCAACCAAGTCACTTCTGAGAATAGTGTATCGCGCGACC 7664
DB 421 CTTTCTCTGACTGGTGTAGTACTCAACCAAGTCACTTCTGAGAATAGTGTATCGCGCGACC 362
QY 7665 GAGTTGCTCTTTCGCGGCGTCAACACGGGATTAATACCGGCGCATACAGCAACTTTAAA 7724
DB 361 GAGTTGCTCTTTCGCGGCGTCAATACGGGATTAATACCGGCGCATACAGCAACTTTAAA 302
QY 7725 AGTGTCTCATCTTGGAAAACGTTCTTCGGGGCGGAAAACCTCTCAAGGATCTTACCGCTGT 7784
DB 301 AGTGTCTCATCTTGGAAAACGTTCTTCGGGGCGGAAAACCTCTCAAGGATCTTACCGCTGT 242
QY 7785 GAGATCCAGTTTCGATGTAAACCACTCGTGTGACCAACTGATCTTCAGCATCTTTTACTTT 7844
DB 241 GAGATCCAGTTTCGATGTAAACCACTCGTGTGACCAACTGATCTTCAGCATCTTTTACTTT 182
QY 7845 CACCAGCGTTTCTGCGTGAGCAAAAACAGGAGGCAAAATCCCGCAAAAAGGGAATAAG 7904
DB 181 CACCAGCGTTTCTGCGTGAGCAAAAACAGGAGGCAAAATCCCGCAAAAAGGGAATAAG 122

```

QY 7905 GCGACACGGAATGTTGAATACATCACTCTCTCTTTCAATATATTATTAAGACATTTA 7964
|
Db 121 GCGACACGGAATGTTGAATACATCACTCTCTCTTTCAATATATTATTAAGACATTTA 62
|
QY 7965 TCAGGTTATTGTTCTCATAGCGGATACATATTTGAATATTGAAATATTGAAAAATAACAAAT 8024
|
Db 61 TCAGGTTATTGTTCTCATAGCGGATACATATTTGAATATTGAAATATTGAAAAATAACAAAT 2
|
QY 8025 A 8025
|
Db 1 A 1

RESULT 9
CD459085/c
LOCUS
DEFINITION
F908_08d03 R F908 AAFEC ECORC Fusarium graminearum complex_substrate
Gibberella zeae cDNA clone F908_08d03, mRNA sequence.
ACCESSION
CD459085
VERSION
CD459085.1 GI:31373825
KEYWORDS
EST.
SOURCE
Gibberella zeae
ORGANISM
Gibberella zeae
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE
1 (bases 1 to 872)
AUTHORS
Watson,R.J., Heyes,R., Chapados,J., Couroux,P., Harris,L.J.,
Hattori,J., Lacroix,C., Ouellet,I., Robert,L.S., Singh,J.A.,
Spott,D. and Tinker,N.A.
TITLE
A cDNA library prepared from Fusarium graminearum grown on a
complex plant substrate
JOURNAL
Unpublished (2003)
COMMENT
Contact: Watson, Robert.J.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-food Canada
Hlqg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
CANADA
Tel: (613) 759-1655
Fax: (613) 759-1701
Email: watsonrj@agr.gc.ca.

FEATURES
Location/Qualifiers
1..872
/organism="Gibberella zeae"
/mol_type="mRNA"
/strain="DAOM 180378"
/db_xref="taxon:5518"
/clone="F908_08d03"
/tissue_type="Mycelium"
/dev stage="Asexual"
/lab_hosts="E. coli DH10B"
/clone_lib="F908_AAFEC_ECORC_Fusarium_graminearum_complex_s
ubstrate"
note=vector: pBluescript II+; Site 1: EcoRI; Site 2:
XhoI; Fusarium graminearum grown on a complex plant
substrate-- wheat leaves treated to remove most of the low
molecular weight, water-soluble components."

ORIGIN
Query Match 10.1%; Score 829.6; DB 14; Length 872;
Best Local Similarity 98.9%; Pred. No. 1.2e-173;
Matches 829; Conservative 5; Mismatches 43; Indels 0; Gaps 0;

QY 7215 CCCACGCTCACCGGCTCCAGATTATCAGCAATAAACACGCGGAGGCGGAGG 7274
|
Db 872 CCCACGCTCACCGGCTCCAGATTATCAGCAATAAACACGCGGAGGCGGAGG 813
|
QY 7275 CAGAGTGGTCCGACCTTATCCGCTCCATCCAGTCTATTATTTGCGGGAGC 7334
|
Db 812 CAGAGTGGTCCGACCTTATCCGCTCCATCCAGTCTATTATTTGCGGGAGC 753
|
QY 7335 TAGAGTAAGTAGTTCGCGCAGTTAATAGTTTGCACACGTTGTTGCCATTGTCGAGGCAT 7394
|
Db 752 TAGAGTAAGTAGTTCGCGCAGTTAATAGTTTGCACACGTTGTTGCCATTGTCGAGGCAT 693

```

```

QY 7395 CGTGGTGTCAAGCTCGTCTGTTGGTATGGCTTCATTTCAGCTCCGTTCCCAACGATCAAG 7454
|
Db 692 CGTGGTGTCAAGCTCGTCTGTTGGTATGGCTTCATTTCAGCTCCGTTCCCAACGATCAAG 633
|
QY 7455 GCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGTTAGTCTCTTCGGTCCGAT 7514
|
Db 632 GCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGTTAGTCTCTTCGGTCCGAT 573
|
QY 7515 CGTTGTCAAGTAAGTTCGCGCAGTGTATTCATCATGTTATGGCAGACATGCATAA 7574
|
Db 572 CGTTGTCAAGTAAGTTCGCGCAGTGTATTCATCATGTTATGGCAGACATGCATAA 513
|
QY 7575 TTCTTCTACTGTTCATCCCATCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCA 7634
|
Db 512 TTCTTCTACTGTTCATCCCATCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCA 453
|
QY 7635 GTCATTCCTGAGATAGTGTATCGCGGACCGAGTTGCTCTTCCCGCGCTCAACACGGA 7694
|
Db 452 GTCATTCCTGAGATAGTGTATCGCGGACCGAGTTGCTCTTCCCGCGCTCAACACGGA 393
|
QY 7695 TAATACCGCGCCACATAGCAGAACTTTAAAGTGTCTATCATTTGGAACCTTTCTTCGG 7754
|
Db 392 TAATACCGCGCCACATAGCAGAACTTTAAAGTGTCTATCATTTGGAACCTTTCTTCGG 333
|
QY 7755 GCGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAAACCCACTCGTGC 7814
|
Db 332 GCGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAAACCCACTCGTGC 273
|
QY 7815 ACCCACTGATCTTCAGCATCTTTTACTTTTACCAGCTTTTCTGGTGAGCAAAAACAGG 7874
|
Db 272 ACCCACTGATCTTCAGCATCTTTTACTTTTACCAGCTTTTCTGGTGAGCAAAAACAGG 213
|
QY 7875 AAGGCAAAATGCGCAAAAAGGAATAGGCGACACGGAATGTTGAAATCTCACTACT 7934
|
Db 212 AAGGCAAAATGCGCAAAAAGGAATAGGCGACACGGAATGTTGAAATCTCACTACT 153
|
QY 7935 CTTCTCTTTTCAATATTATTGAAGCATTTATCAGGTTTATTGTCTCATGAGCGGATACAT 7994
|
Db 152 CTTCTCTTTTCAATATTATTGAAGCATTTATCAGGTTTATTGTCTCATGAGCGGATACAT 93
|
QY 7995 ATTGTAATGTTATGAAAAATAACAAATAGGGTTCGCGCAGCATTTCCCGGAAAA 8052
|
Db 92 ATTGTAATGTTATGAAAAATAACAAATAGGGTTCGCGCAGCATTTCCCGGAAAA 35

RESULT 10
BM438950/c
LOCUS
DEFINITION
IctLvr00491 Liver cDNA library Ictalurus punctatus cDNA 5', mRNA
Sequence.
ACCESSION
BM438950
VERSION
BM438950.1 GI:18460672
KEYWORDS
EST.
SOURCE
Ictalurus punctatus (channel catfish)
ORGANISM
Ictalurus punctatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
Ictaluridae; Ictalurus.
1 (bases 1 to 854)
Feng,J., Kucuktas,H., Kocabas,A., Li,P. and Liu,Z.
Transcriptome of channel catfish (Ictalurus punctatus): initial
analysis of expressed sequence tags from the liver
Unpublished (2002)
Contact: Liu ZJ
The Fish Molecular Genetics and Biotechnology Laboratory,
Department of Fisheries and Allied Aquacultures and Program of Cell
and Molecular Biosciences
Auburn University
203 Swingle Hall, Auburn University, Auburn, AL 36849, USA
Tel: 334 844 4054
Fax: 334 844 9208
Email: zliu@acesag.auburn.edu

```


Db	15	AGGAGGCAAAATGC	1
RESULT 11			
CD458333/c			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
1. .870			
/organism="Gibberella zeae"			
/mol_type="mRNA"			
/strain="DAOM 180378"			
/db_xref="taxon:5518"			
/clone="Fg08_09e07"			
/tissue_type="Mycelium"			
/dev_stage="Asexual"			
/lab_host="E. coli DH10B"			
/clone_lib="Fg08_AAFc_ECORC_Fusarium_graminearum_complex_s"			
/note="Vector: pBluescript II+; Site_1: EcoRI; Site_2: XhoI; Fusarium graminearum grown on a complex plant substrate-- wheat leaves treated to remove most of the low molecular weight, water-soluble components."			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
9.9%; Score 819.2; DB 14; Length 870;			
98.7%; Pred. No. 2.6e-171;			
843; Conservative			
3; Mismatches			
6; Indels			
2; Gaps			
2; 7219			
CGCTCACCGGCTCCGAGTTATCAGCAATAAACCCAGCCAGCGGAGCGCGAGCGAGA			
7278			
870			
CGCTCWCAGGCTCCGAGTTATCAGCAATAAACCCAGCCAGCGGAGCGCGAGCGAGA			
811			
Db			
7279			
AGTGTCCTGCACTTTATCCGCTCCATCGAGTCTATTATTGTTGCCGGAAGCTAGA			
7338			
810			
AGTGTCCTGCACTTTATCCGCTCCATCGAGTCTATTATTGTTGCCGGAAGCTAGA			
751			
QY			
7339			
GTAAGTAGTTCGCACTTAATAGTTTGGCAACGTTTGTTCATTGCTGTCAGGATCGTG			
7398			
Db			
750			
GTAAGTAGTTCGCACTTAATAGTTTGGCAACGTTTGTTCATTGCTGTCAGGATCGTG			
691			
QY			
7399			
GTGTCAAGTTCGTTGGTATGCTTCATTTCAGCTCCGTTCCCAACGATCAAGCGGA			
7458			
Db			
690			
GTGTCAAGTTCGTTGGTATGCTTCATTTCAGCTCCGTTCCCAACGATCAAGCGGA			
631			
QY			
7459			
GTTACATGATCCCGATGTTGTGCAAAAAGCGTTAGCTCCTTCGTTCCCGATCGTT			
7518			
Db			
630			
GTTACATGATCCCGATGTTGTGCAAAAAGCGTTAGCTCCTTCGTTCCCGATCGTT			
572			
QY			
7519			
GTGAG-AGGTAAGTTGGCGGAGCTGTTATCACTCATGTTATGGCAGCTGCATAATTC			
7577			

Db	15	AGGAAGGCAAAATGC 1
RESULT 11		
CD458333/c		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
1. .870		
/organism="Gibberella zeae"		
/mol_type="mRNA"		
/strain="DAOM 180378"		
/db_xref="taxon:5518"		
/clone="Fg08_09e07"		
/tissue_type="mycelium"		
/dev_stage="Asexual"		
/lab_host="E. coli DH10B"		
/clone_lib="Fg08_AAFc_EORC_Fusarium_graminearum_complex_s"		
/note="Vector: pBluescript II+; Site 1: EcoRI; Site 2: XhoI; Fusarium graminearum grown on a complex plant substrate-- wheat leaves treated to remove most of the low molecular weight, water-soluble components."		
ORIGIN		
Query Match		
Best Local Similarity 9.9%; Score 819.2; DB 14; Length 870;		
Matches 843; Conservative 3; Mismatches 6; Indels 2; Gaps 2;		
QY	7219	CGCTCACCGGCTCCAGATTATCAGCAATAAACCCAGCCAGCCGGAAGCGCGCAGA 7278
Db	870	CGCTCMCGGCTCCAGATTATCAGCAATAAACCCAGCCAGCCGGAAGCGCGCAGA 811
QY	7279	AGTGCTCTGCAACTTTATCCGCTCCATCCAGTCTATTATTTGTCGGGAAGCTAGA 7338
Db	810	AGTGCTCTGCAACTTTATCCGCTCCATCCAGTCTATTATTTGTCGGGAAGCTAGA 751
QY	7339	GTAAGTAGTTCGCCAGTTAATAGTTTGGCAACCTTGTTCGCAATGCTACAGGCATCGT 7398
Db	750	GTAAGTAGTTCGCCAGTTAATAGTTTGGCAACCTTGTTCGCAATGCTACAGGCATCGT 691
QY	7399	GTGTACAGCTCGTGTGGTATGGCTTCAATCAGCTCCGTTCCCAACGATCAAGCGGA 7458
Db	690	GTGTACAGCTCGTGTGGTATGGCTTCAATCAGCTCCGTTCCCAACGATCAAGCGGA 631
QY	7459	GTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCTCCCATCGTT 7518
Db	630	GTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCTCCCATCGTT 572
QY	7519	GTCAG-AGTAAAGTGGCCGAGTGTATATCTCATGGTTATGGCAGCACTGCATAATTC 7577

Qy	7940	TTTTTCAATATATTGAAGCATTTATCAGGGTTATTCTCTCATGAGCGGATACATATTG	7999
Db	170	TTTTTCAATATATTGAAGCATTTATCAGGGTTATTCTCTCATGAGCGGATACATATTG	111
Qy	8000	AATGTATTAGAAAAATA	8017
Db	110	AATGTATTAGAAAAATA	93

RESULT 13
CD649375/c
LOCUS
DEFINITION CD649375 819 bp mRNA linear EST 18-JUN-2003
CD649375 Crassostrea virginica Gonad Crassostrea virginica cDNA
5', mRNA sequence.
ACCESSION
CD649375
CD649375.1 GI:31906346
KEYWORDS
EST.

SOURCE	Crassostrea virginica (eastern oyster)
ORGANISM	Crassostrea virginica Eukarya; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Ostreoida; Ostreidae; Ostreidae; Crassostrea. 1 (bases 1 to 819)
REFERENCE	Peatman,B., Kucuktas,H., Li,P., He,C., Feng,J., Wei,X. and Liu,Z.
AUTHORS	Differentially expressed oyster (Crassostrea virginica) genes after
TITLE	exposure to mercury
JOURNAL	Unpublished (2003)

CONTACT: Liu Zhi
The Fish Molecular Genetics and Biotechnology Laboratory,
Department of Fisheries and Allied Aquacultures and Program of Cell
and Molecular Biosciences
Auburn University
203 Swingle Hall, Auburn University, Auburn, AL 36849, USA
Tel: 334 844 4054
Fax: 334 844 9208
Email: zliu@aesag.auburn.edu

```

FEATURES
source
1. 819
/organism="Crassostrea virginica"
/mol_type="mRNA"
/db_xref="taxon:6565"
/clone_lib="Crassostrea virginica Gonad"
/note="Organ: Gonad; Vector: pSport1; Site_1: NotI;
Site_2: SalI"

```

[illegible]

DB	819	ACAGTTACCAATGCTTAAATCAGTGGAGCACCCTATCTCAGCGATCTGCTCTATTTCGTTTCAT	760
QY	7127	CCATAGTTGCTGACTCCCGCGTGGTGGAGATAACTACGATACGGGGGGGCTTACCATCTG	7186
DB	759	CCATAGTTGCTGACTCCCGCGTGGTGGAGATAACTACGATACGGGGGGCTTACCATCTG	700
QY	7187	GCCCCAGTGTGCAATGATACCGGGAGACCCAGCTCACCGCTCCAGATTTATCAGCAA	7246
DB	699	GCCCCAGTGTGCAATGATACCGGGAGACCCAGCTCACCGGTTCAGATATCAGCAA	640

Qy	7247	TAAACGAGCAGCGGGAAGGCGGAGCGCAGAGAGTGCTCTGCAACTTTTATCGGCTTCCA	7306
Db	639	TAAACGAGCAGCGGGAAGGCGGAGCGCAGAGAGTGCTCTGCAACTTTTATCGGCTTCCA	580
Qy	7307	TCGAGTCTATTATGTTGCGGGAAGCTAGAGTATAGTATTCGCCAGTTAAAGATTGC	7366
Db	579	TCGAGTCTATTATGTTGCGGGAAGCTAGAGTATAGTATTCGCCAGTTAAAGATTGC	520

Qy 7367 GCAACGTTGTGGCATTGCTGCAGGCATCGTGGTGTCACGCTCGTCGTTTGGGTATGGCTT 742

Dd 519 GCAACGTTGTGGCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGGTATGGCTT 460

QY	7427	CATTCACTCCGGTTCCCAACGATCAAGCGAGTTACATGATCCCCCATGTTGTGCAAAA	74886
Db	459	CATTCACTCCGGTTCCCAACGATCAAGCGAGTTACATGATCCCCCATGTTGTGCAAAA	400
QY	7487	AAGCGGTTAGTCTCTTCGGTCCCTCCGATCGTTGTCAGAAGTAAAGTTGGCGCGAGTGTAT	7546
Db	399	AAGCGGTTAGCT-CTTCGGTCCCTCCGATCGTTGTCAGAAGTAAAGTTGGCGCGAGTGTAT	341
QY	7547	CACATCATGTTATGGGAGCACTGCATAAATCTCTTACTGTCATGCCATCCGTAAGATGCT	7606
Db	340	CACATCATGTTATGGGAGCACTGCATAAATCTCTTACTGTCATGCCATCCGTAAGATGCT	281
QY	7607	TTTCTGTGACTGTGTGAGTACTCAACCAAGATCAATCTGAGAATAGTGTATGCGGCGACCGA	7666
Db	280	TTTCTGTGACTGTGTGAGTACTCAACCAAGATCAATCTGAGAATAGTGTATGCGGCGACCGA	221
QY	7667	GTTGCTCTTGCCGGCGTCAACACGGGATAAATACCGGCCACATACAGAACTTTAAAG	7726
Db	220	GTTGCTCTTGCCGGCGTCAANTACGGGATAAATACCGGCCACATACAGAACTTTAAAG	161
QY	7727	TGCTCATCTATGGAAAAAGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGA	7786
Db	160	TGCTCATCTATGGAAAAAGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGA	101
QY	7787	GATCCAGTTCGATGTAAACCCACTCGTGCAACCACTGATCTTCAGCATCTTTTACTTTCA	7846
Db	100	GATCCAGTTGATGTAAACCACTCGTGCAACCACTGATCTTCAGCATCTTTTACTTTCA	41
QY	7847	CCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGC	7886
Db	40	CCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGC	1

RESULT 14
B2576726/c
LOCUS
DEFINITION msh2_5071.y2 msh pseudomonas aeruginosa genomic clone msh2_5071,
1011 bp DNA linear GSS 17-DEC-2002

ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
B2576726	B2576726.1	GSI.	Pseudomonas aeruginosa	Pseudomonas aeruginosa
B2576726.1	GI:27211787	GSI.	Pseudomonas aeruginosa	Pseudomonas aeruginosa
			Bacteria; Proteobacteria;	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
			Pseudomonadaceae; Pseudomonas.	Pseudomonadaceae; Pseudomonas.

1 (bases 1 to 1011)
 Spencer D.H., Raymond C.K., Smith E.E., Hastings M.,
 Burns J.L., Kaul R. and Olsen M.V.
 Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
 J. Bacteriol. (2002) in press
 Contact: Chris K. Raymond
 Genome Center
 University of Washington

Box 352145, Seattle, WA 98105-2145, USA
Tel: 20622216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: Shotgun.
Location/Qualifiers
1. 1011
/organism="Pseudomonas aeruginosa"
/mol type="genomic DNA"

```

/strain="msh"
/db_xref="taxon:287"
/clone="msh2_5071"
/clone_lib="msh"
/notes="Environmental isolate. Whole genomic shotgun
library."

ORIGIN
Query Match          9.6%   Score 794.4;   DB 28;   Length 1011;

```

Best Local Similarity 93.7%; Pred. No. 8.8e-166;
Matches 869; Conservative 0; Mismatches 54; Indels 4; Gaps 4;

QY 7017 AAAATGAAGTTTAAATCAATCACTAACTATATATAGTAACCTGCTCAGAGTTACCA 7076
Db 993 AAAATGAAGTTTAAATCAATCACTAACTATATAGTAACCTGCTCAGAGTTACCA 935

QY 7077 ATGCTTAATCACTGAGGAGCCATCTCTCAGGAGTCT-GTCTATTTCCTTCATCCATAGTTG 7135
Db 934 ATGCTTAATCACTGAGGAGCCATCTCTCAGGAGTCT-GTCTATTTCCTTCATCCATAGTTG 875

QY 7136 CCGTACTCCCGTCTGTAGTAACTACGATAGGAGGCTTACCATCTGGCCCGAGTG 7195
Db 874 CCGTACTCCCGTCTGTAGTAACTACGATAGGAGGCTTACCATCTGGCCCGAGTG 815

QY 7196 CTGCAATGATACCGGAGACCCAGCTT-CACCGGCTCCAGATTTATCAGCAATAAACCCAG 7254
Db 814 TTGCAATGATACCGGAGACCCAGCTT-CACCGGCTCCAGATTTATCAGCAATAAACCCAG 755

QY 7255 COAGCGGAGGCGGAGCGAGAGTGTCTGCACTTTATCCGCTCCATCCAGTCT 7314
Db 754 COAGCGGAGGCGGAGCGAGAGTGTCTGCACTTTATCCGCTCCATCCAGTCT 695

QY 7315 ATTAATTTCCCGGAGCTAGATAGTGTGCGCACTTAATAGTTTGGCAACGTT 7374
Db 694 ATTAATTTCCCGGAGCTAGATAGTGTGCGCACTTAATAGTTTGGCAACGTT 636

QY 7375 GTTGCCATTGCTGAGGAGCATGCTGCTCAGCTCGTGTGCTGATGCTTCACTCAG 7434
Db 635 GTTGCCATTGCTGAGGAGCATGCTGCTCAGCTCGTGTGCTGATGCTTCACTCAG 576

QY 7435 TCCGCTTCCCAACGATCAAGCGAGTTACATGATCCCGCATGTTGTCAGAAAGCGTT 7494
Db 575 TCCGCTTCCCAACGATCAAGCGAGTTACATGATCCCGCATGTTGTCAGAAAGCGTT 516

QY 7495 AGCTCTCTCGTCTCCGATCGTGTGTCAGAGTAAGTTGGCCGAGTGTATCACTCAG 7554
Db 515 AGCTCTCTCGTCTCCGATCGTGTGTCAGAGTAAGTTGGCCGAGTGTATCACTCAG 456

QY 7555 GTTATGGCAGCACTGCATAATTTCTTCTTCTGATGCTATCCGATCCGTAAGATGCTTTCTG 7614
Db 455 GTTATGGCAGCACTGCATAATTTCTTCTTCTGATGCTATCCGATCCGTAAGATGCTTTCTG 396

QY 7615 ACTGGTGAATCAACCACTGCTGAGATGCTGAGATGCTGATGCGGAGCGAGTTGCTCT 7674
Db 395 ACTGGTGAATCAACCACTGCTGAGATGCTGAGATGCTGATGCGGAGCGAGTTGCTCT 336

QY 7675 TGCCCGGCTCAACACCGGATAATACCGGCCACATAGCAGAACTTTAAAGTGCTCATC 7734
Db 335 TGCCCGGCTCAACACCGGATAATACCGGCCACATAGCAGAACTTTAAAGTGCTCATC 276

QY 7735 ATTGGAACCTCTTCCGGGCGAATCTCAAGATCTTCAAGATCTTCAAGATCTTCAAG 7794
Db 275 ATTGGAACCTCTTCCGGGCGAATCTCTCAAGATCTTCAAGATCTTCAAGATCTTCAAG 216

QY 7795 TCGATGAACCACTCTGTCACCACTGATCTTCAAGATCTTCAAGATCTTCAAGATCTTCAAG 7854
Db 215 TCGATGAACCACTCTGTCACCACTGATCTTCAAGATCTTCAAGATCTTCAAGATCTTCAAG 156

QY 7855 TCTGGTGAACCAACAGAGGCAAAATGCGCAAAAGGATAGGCGGACAGG 7914
Db 155 TCTGGTGAACCAACAGAGGCAAAATGCGCAAAAGGATAGGCGGACAGG 96

QY 7915 AAATGTTGAATCACTATCTTCTCTT 7941
Db 95 AAATGTTGAATCACTATCTTCTCTT 69

RESULT 15

BZ572566

LOCUS

DEFINITION BZ572566 1574 bp DNA linear GSS 17-DEC-2002
msh2_2693.x1 msh Pseudomonas aeruginosa genomic clone msh2_2693,
genomic survey sequence.

ACCESSION BZ572566
VERSION BZ572566.1 GI:27207627
KEYWORDS GSS
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
REFERENCE Pseudomonas aeruginosa
AUTHORS Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
1 (bases 1 to 1574)
Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA.
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
Location/Qualifiers
1. .1574
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="msh"
/db_xref="taxon:287"
/clone="msh2_2693"
/clone_lib="msh"
/notes="Environmental isolate. Whole genomic shotgun
library."

ORIGIN

Query Match 9.5%; Score 785.4; DB 28; Length 1574;
Best Local Similarity 92.9%; Pred. No. 9.5e-164;
Matches 903; Conservative 0; Mismatches 54; Indels 15; Gaps 7;

QY 6283 AACGCTAAAGAGCCGCTGCTGCGTTTTCATAGGCTCCGCCCCCTGACGAGCAT 6342
Db 91 AACGCTAAAGAGCCGCTGCTGCGTTTTCATAGGCTCCGCCCCCTGACGAGCAT 150

QY 6343 CACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACCAG 6402
Db 151 CACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACCAG 210

QY 6403 GCCTTTCCCGCTGGAAAGCTCCCTCGTGGCTCTCTCTGTTCCGACCCCTGCGCTTACCGGA 6462
Db 211 GCCTTTCCCGCTGGAAAGCTCCCTCGTGGCTCTCTCTGTTCCGACCCCTGCGCTTACCGGA 270

QY 6463 TACTGTGCGCTTTCTCCCTTCGGGAGCGTGGGCTTCTCATAGCTCAGCTGTAGG 6522
Db 271 TACTGTGCGCTTTCTCCCTTCGGGAGCGTGGGCTTCTCATAGCTCAGCTGTAGG 330

QY 6523 TATCTCAGTTCGGGTAGGTCTTCCCTCCAAAGCTGGGCTGTGTGCAAGAACCCCGCTT 6582
Db 331 TATCTCAGTTCGGGTAGGTCTTCCCTCCAAAGCTGGGCTGTGTGCAAGAACCCCGCTT 390

QY 6583 CAGCCGACCGCTGCGCTTATCCCGTAACCTATCGCTTGTAGTCCCAACCCCGTAAGACAC 6642
Db 391 CAGCCGACCGCTGCGCTTATCCCGTAACCTATCGCTTGTAGTCCCAACCCCGTAAGACAC 450

QY 6643 GACTTATCGGCTAGGAGCGCTCTGCTGAGGATAGGAGGAGGATAGTATAGG 6702
Db 451 GACTTATCGGCTAGGAGCGCTCTGCTGAGGATAGGAGGAGGATAGTATAGG 510

QY 6703 GGTGCTACAGAGTCTTGAAGTGGTGGCTTAACCTACGCTACACTAGAAGACAGTATTT 6762
Db 511 GGTGCTACAGAGTCTTGAAGTGGTGGCTTAACCTACGCTACACTAGAAGACAGTATTT 570

QY 6763 GGTATCTGCGCTCTGCTGAAGCCAGTATACCTTCGGAAGAGAGTGTGTAGCTCTGTATCC 6822
Db 571 GGTATCTGCGCTCTGCTGAAGCCAGTATACCTTCGGAAGAGAGTGTGTAGCTCTGTATCC 630

Qy	6823	GGCAACAAACCAACCGCTGTAGCGTGG-TTTTTTTGTTTGCAAGCAGCAGATTACGCG	6881
Db	631	GGCAACAAACCAACCGCTGTAGCGTGGTTTTTTTGTGCAAGCAGCAGATTACGCG	690
Qy	6882	CAGAAAAAGGATCTCAGAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTG	6941
Db	691	CAGAAAAAGGATCTCAGAGATCCTTTGATC-TTCTACGGGTCTGACGCTCAGTG	749
Qy	6942	GAACGAAACTCAGGTTAAGGATTTTGGTCATGAGATTATCAAAAAAGGATCTTCACCTA	7001
Db	750	GAACGAAACTCAGGTTAAGGGA-TTTGGTCATGAGATTATCAAAAAAGGATCTTCACCTA	808
Qy	7002	GATCCTTTTAAATTAATAAGTATTAATCAATCTAAAGTATATGAGTAAACTTG	7061
Db	809	GATCCTTTTAAATTAATAAGTATTAAG-TTTAAATAATCTAAAGTATATGAGTAAACTTG	867
Qy	7062	GTCTGACAGTTACCAATGCTTAATCAGTGGGACCTATCTCAGCGATCTGTATTTCG	7121
Db	868	GTCTGACAGTACCAATGGCTATCAGTGGGACCTATCTTACCGATCTGTATTTCG	927
Qy	7122	TTTATCCATAGTTGCCTGACTCCCGGTCGTGTAGATAACTACGATACGGAGGGCTTACC	7181
Db	928	-TCATCCATAGTTGGCGGACTCCCGGGTAAAAACAAC-----AATCCGAGGGCTTAC	981
Qy	7182	ATCTGGCCCCAGTGTCTCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTATC	7241
Db	982	ATTGGGCCCCAGTGTGAAATACCGGAGACCCACCTCCCGG-----TCCGATTATC	1036
Qy	7242	AGCAATAAAC	7251
Db	1037	AGCAATAACC	1046

Search completed: September 24, 2004, 09:58:30
Job time : 12509.4 secs

This Page Blank (copy)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2004, 02:36:16 ; Search time 1874.44 Seconds
(without alignments)
18677.278 Million cell updates/sec

Title: US-10-030-390-3

Perfect score: 8241
Sequence: 1 agatatacatccaagacg.....aaaataacagtattattcg 8241

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001s:*
5: Geneseqn2002s:*
6: Geneseqn2003as:*
7: Geneseqn2003bs:*
8: Geneseqn2003cs:*
9: Geneseqn2003ds:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8241	100.0	8241	5	Aaf29647 Plasmid p
2	7555.6	91.7	8598	2	Aaf29647 Plasmid p
3	7260	88.1	8974	2	Aaf29647 Plasmid p
4	6687.6	81.2	9408	4	Aaf29647 Plasmid p
5	6675.6	81.0	9407	4	Aaf29647 Plasmid p
6	6582	79.9	9191	2	Aaf29647 Plasmid p
7	6561.6	79.6	9276	4	Aaf29647 Plasmid p
8	6561.6	79.6	9276	4	Aaf29647 Plasmid p
9	6561.6	79.6	9276	7	Aaf29647 Plasmid p
10	6561.6	79.6	9276	7	Aaf29647 Plasmid p
11	6561.6	79.6	9276	7	Aaf29647 Plasmid p
12	6561.6	79.6	9276	7	Aaf29647 Plasmid p
13	6561.6	79.6	9276	9	Aaf29647 Plasmid p
14	6561.6	79.6	9276	9	Aaf29647 Plasmid p
15	6383.8	77.5	9776	6	Aaf29647 Plasmid p
16	6383.8	77.5	9776	6	Aaf29647 Plasmid p
17	6009	72.9	10215	2	Aaf29647 Plasmid p
18	5709.8	69.3	10462	4	Aaf29647 Plasmid p
19	5709.8	69.3	10462	6	Aaf29647 Plasmid p
20	5677.8	68.9	10492	4	Aaf29647 Plasmid p
21	5665.8	68.8	10491	6	Aaf29647 Plasmid p
22	5657.8	68.7	10512	6	Aaf29647 Plasmid p
23	5657.8	68.7	10512	6	Aaf29647 Plasmid p

24	5142.8	62.4	8584	2	AAT35166
25	5124.8	62.2	8590	2	AAT35168
26	4734.6	57.5	9156	2	AAT77817
27	4555.2	55.4	8137	2	AAT35165
28	2657.4	32.2	2659	2	AAT35165
29	2404.6	29.2	5740	7	ACC42391
30	2312.4	28.1	11466	6	ABQ74937
31	2312.4	28.1	11466	6	ABQ74937
32	2312.4	28.1	11466	6	ABQ74937
33	2312.4	28.1	11466	6	ABQ74937
34	2312.4	28.1	11466	6	ABQ74937
35	2312.4	28.1	11466	6	ABQ74937
36	2309.4	28.0	19798	4	AAF83671
37	2309.4	28.0	19798	4	AAF83671
38	2309.4	28.0	20160	4	AAF83672
39	2309.4	28.0	20160	4	AAF83672
40	2309.4	28.0	20247	4	AAF83675
41	2309.4	28.0	20247	4	AAF83675
42	2309.2	28.0	10111	6	AAH99982
43	2306.8	28.0	3621	1	AAH99982
44	2306	28.0	34303	2	AAV07261
45	2305.2	28.0	3547	1	AAH99982

ALIGNMENTS

RESULT 1
AAF29647
ID AAF29647 standard; DNA; 8241 BP.
XX
XX AAF29647;
XX
XX AAF29647;
DT 10-APR-2001 (first entry)
XX
DE Plasmid pPICMURF1.
XX
XX Mouse TFF1; mTFF; trefoil protein; antiinflammatory; gastrointestinal;
XX anti-ulcer; peptide therapy; gastrointestinal disease; acute colitis;
XX Crohn's disease; ulcerative colitis; plasmid pPICMURF1;
XX recombinant vector; ds.
XX Synthetic.
XX
XX WO200102570-A1.
XX
PD 11-JAN-2001.
XX
XX 05-JUL-2000; 2000WO-EP06343.
XX
XX 05-JUL-1999; 99EP-00870143.
XX
XX (VIAA-) VIAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
XX Hans WC, Steidler L, Remaut ER;
XX
XX WPI; 2001-138142/14.
XX
XX Recombinant Lactococcus lactis for delivering a trefoil peptide useful
XX for treating acute or chronic gastrointestinal inflammatory diseases or
XX disorders, e.g. acute or ulcerative colitis, acute flare-ups of Crohn's
XX disease.
XX
XX Claim 18; Fig 1c; 44pp; English.
XX
XX The present sequence is a recombinant vector which may be used to
XX generate a recombinant Lactococcus lactis capable of delivering a trefoil
XX peptide in vivo. The recombinant microorganism is useful for
XX manufacturing an agent for the delivery of a trefoil peptide to the
XX gastrointestinal tract, and for treating gastric or intestinal diseases
XX or disorders, or lesions caused by gastric or intestinal diseases or
XX disorders. The microorganism may also be used for preparing medicament to
XX be used for treating gastric and/or gastrointestinal diseases or

disorders, acute gastrointestinal inflammatory diseases (e.g., acute colitis, acute flare-ups of Crohn's disease, or ulcerative colitis), and chronic and spontaneously recurring diseases of the gastrointestinal tract comprising Crohn's disease (enteritis regionalis) and ulcerative colitis (colitis ulcerosa). Disease states which can be treated by the method or compositions comprising the recombinant macroorganism or trefoll peptides include disorders of and damage to the alimentary canal, including the mouth, oesophagus, stomach and large and small intestine, as well as for the protection and treatment of tissues that lie outside the alimentary canal.

Sequence 8241 BP: 2193 A: 1963 C: 1870 G: 2215 T: 0 U: 0 Other: 0

Query Match 100.0%: Score 8241: DB 5: Length 8241:

Query Match	100.0%;	Score 8241;	DB 5;	length 8241;
Best Local Similarity	100.0%;	Pred. No. 0;		

BEST LOCAL SIMILARITY 100.0%; PRED. NO. 0;
Matches 8241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AGATCTAACATCCAAAGACGAAAGCTTGAATGAAACCTTTTGGCAATCCGACATCCACAG	60
Db	1	AGATCTTAACATCCAAAGACGAAAGCTTGAATGAAACCTTTTGGCAATCCGACATCCACAG	60
QY	61	GTCCATTCTCACACATAGTGCCTAAACGCAACGAGGAGGATACACTAGCACGACGCGT	120
Db	61	GTCCATTCTCACACATAGTGCCTAAACGCAACGAGGAGGATACACTAGCACGACGCGT	120
QY	121	TGCAAAACGAGGACCTCCACTCTCTCTCTCAACACCCACCTTTTGGCATCGAAAAACC	180
Db	121	TGCAAAACGAGGACCTCCACTCTCTCTCTCAACACCCACCTTTTGGCATCGAAAAACC	180
QY	181	AGCCAGATTATTTGGCTTGATGGAGTCGCTTCAATTCCTTCTATTAGGTACTA	240
Db	181	AGCCAGATTATTTGGCTTGATGGAGTCGCTTCAATTCCTTCTATTAGGTACTA	240
QY	241	ACACCATGACTTTATTAGCCCTGTCTATCTGGCCCCCTGGCGAGGTCATGTTTGTTTA	300
Db	241	ACACCATGACTTTATTAGCCCTGTCTATCTGGCCCCCTGGCGAGGTCATGTTTGTTTA	300
QY	301	TTTCCGAATGCAACAAAGCTCCGCATTACCCGAAACATCACTCCAGATGAGGCTTTCTG	360
Db	301	TTTCCGAATGCAACAAAGCTCCGCATTACCCGAAACATCACTCCAGATGAGGCTTTCTG	360
QY	361	AGTGTGGGTCAAAATAGTTTCATGTTCCCAAAATGCCCCAAACCTGACAGTTTAAACGCT	420
Db	361	AGTGTGGGTCAAAATAGTTTCATGTTCCCAAAATGCCCCAAACCTGACAGTTTAAACGCT	420
QY	421	GTCTTGGAACTTAATATGACAAAAGCGTGATCTCATCCAAAGATGAACCTTGGTTTCG	480
Db	421	GTCTTGGAACTTAATATGACAAAAGCGTGATCTCATCCAAAGATGAACCTTGGTTTCG	480
QY	481	TTGAAATGCTAACCGCGCAGTTGGTCAAAAAGAAACTTCCAAAAGTCGCGCATACCGTTTGT	540
Db	481	TTGAAATGCTAACCGCGCAGTTGGTCAAAAAGAAACTTCCAAAAGTCGCGCATACCGTTTGT	540
QY	541	CTTGTCTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAATGCTTAGCGCAGTCT	600
Db	541	CTTGTCTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAATGCTTAGCGCAGTCT	600
QY	601	CTCTATCGCTTCTGAAACCCCGGTGACCTGTGCGAAACGCAAAATGGGAAACACCCGCT	660
Db	601	CTCTATCGCTTCTGAAACCCCGGTGACCTGTGCGAAACGCAAAATGGGAAACACCCGCT	660
QY	661	TTTTTGGATGATTATGCAATGTCTCCCATTTGATGCTTCCAGATTCTGGTGGGAATACT	720
Db	661	TTTTTGGATGATTATGCAATGTCTCCCATTTGATGCTTCCAGATTCTGGTGGGAATACT	720
QY	721	GCTGATAGCCTAAAGTTTCATGATCAAAATTTAACTGTTCTTAACCCCTACTTTCAGACGCAAT	780
Db	721	GCTGATAGCCTAAAGTTTCATGATCAAAATTTAACTGTTCTTAACCCCTACTTTCAGACGCAAT	780
QY	781	ATATAAACAGAGAAGAGTGCCTTGCTTAAACCTTTTTTTTATCATATTATAGCTT	840
Db	781	ATATAAACAGAGAAGAGTGCCTTGCTTAAACCTTTTTTTTATCATCATATTATAGCTT	840

1921 CTTGGTATGCGGTACTGCGGGGCTCTTGGGGATATCGTCCATTCGGACAGATCGC 1980
1981 CAGTCACTATGCGGTGCTGCTAGCGCTATATGCGGTGATGCAATTCCTATGCGCACCGGT 2040
1981 CAGTCACTATGCGGTGCTGCTAGCGCTATATGCGGTGATGCAATTCCTATGCGCACCGGT 2040
2041 TCTCGGAGCACTGTCGGACCGCTTTGGCGCGCGCCAGTCTGCTGCTTGGCTTACTTGG 2100
2041 TCTCGGAGCACTGTCGGACCGCTTTGGCGCGCGCCAGTCTGCTGCTTGGCTTACTTGG 2100
2101 AGCCACTATGCACTACGGGATCATGGCGACACACCCCGTCTGCTGATCTATCGAATCTA 2160
2101 AGCCACTATGCACTACGGGATCATGGCGACACACCCCGTCTGCTGATCTATCGAATCTA 2160
2161 AATGTAAGTAAATCTCTAATAATTAATAATGCTCCAGTCTTCCATAGCAACCTTAA 2220
2161 AATGTAAGTAAATCTCTAATAATTAATAATGCTCCAGTCTTCCATAGCAACCTTAA 2220
2221 CAGCATTCGGGTAGCATCTAGACCTTCAACAGCAGCCAGATCCATCACTGCTTGGCCAA 2280
2221 CAGCATTCGGGTAGCATCTAGACCTTCAACAGCAGCCAGATCCATCACTGCTTGGCCAA 2280
2281 TATGTTTCAGTCCCTCAGAGGTAGCTCTTGTGAAGTGAATCACTTCTGGAAGGTTGCAG 2340
2281 TATGTTTCAGTCCCTCAGAGGTAGCTCTTGTGAAGTGAATCACTTCTGGAAGGTTGCAG 2340
2341 TGTAACTCCGCTGATTAAGCGGGCATATCCGTAGCTTGGCAAGTGGTGTGATACCGG 2400
2341 TGTAACTCCGCTGATTAAGCGGGCATATCCGTAGCTTGGCAAGTGGTGTGATACCGG 2400
2401 AGGAGTAATCTCCAACTCTCTGAGAGTAGGCAACCAACAAACAGATCCACGGTGT 2460
2401 AGGAGTAATCTCCAACTCTCTGAGAGTAGGCAACCAACAAACAGATCCACGGTGT 2460
2461 GTACTTGATCAACATPAGAAGACATCTCGATTTGCAGATCAAGTGTTCAGAGCGGT 2520
2461 GTACTTGATCAACATPAGAAGACATCTCGATTTGCAGATCAAGTGTTCAGAGCGGT 2520
2521 ACTGATGGACATTTCCAAAGCCTGCTGATAGTTGCAACCGATAGGTTGTAGAGTGTG 2580
2521 ACTGATGGACATTTCCAAAGCCTGCTGATAGTTGCAACCGATAGGTTGTAGAGTGTG 2580
2581 CAATACACTTGGTACATTTCAACCTTGGCAACTGCAAGCTTGGTGTGAAACAGCAT 2640
2581 CAATACACTTGGTACATTTCAACCTTGGCAACTGCAAGCTTGGTGTGAAACAGCAT 2640
2641 CTTCAATTTCTGGCAAGCTTCTGTGCTATATCGACAGCCAAACAGATCACTTGGGAAT 2700
2641 CTTCAATTTCTGGCAAGCTTCTGTGCTATATCGACAGCCAAACAGATCACTTGGGAAT 2700
2701 CAATACCAATGTTAGCTTGACAGAGGTCTGAGGCAACGAATCTGGATCAGCGTATT 2760
2701 CAATACCAATGTTAGCTTGACAGAGGTCTGAGGCAACGAATCTGGATCAGCGTATT 2760
2761 TATCAGCAATTAACCTAGAACTTCAAGGCCAGAGGCAAGTGTCAATCTACACAGGCGTG 2820
2761 TATCAGCAATTAACCTAGAACTTCAAGGCCAGAGGCAAGTGTCAATCTACACAGGCGTG 2820
2821 ATGTGCTAATTTGAACCACTATCTTGGCAGCAGTAACGAATCTTCTGGACCAATA 2880
2821 ATGTGCTAATTTGAACCACTATCTTGGCAGCAGTAACGAATCTTCTGGACCAATA 2880
2881 TTTTGTACACTTAGGACAGTCTGTTCCGTAGCCATAGCAGCTACTGCTTGGGCGC 2940
2881 TTTTGTACACTTAGGACAGTCTGTTCCGTAGCCATAGCAGCTACTGCTTGGGCGC 2940
2941 CTCCTGCTAGCAGATACACTTAGCACCACCTTGTGGGCAACGTAGATGACTTCTGGGG 3000
2941 CTCCTGCTAGCAGATACACTTAGCACCACCTTGTGGGCAACGTAGATGACTTCTGGGG 3000
3001 TAAGGTAACCACTCTTCTTAGGTGAGTAGGCAAAAACAAATTTCTTTCGACCAACCACTT 3060

3001 TAAGGTAACCACTCTTCTTAGGTGAGTAGGCAAAAACAAATTTCTTTCGACCAACCACTT 3060
3061 TGGCAGAAACACCCAGCATCAGGAAAGTGGAAAGGCAAGATTCGCGTTCCACACGGAATAT 3120
3061 TGGCAGAAACACCCAGCATCAGGAAAGTGGAAAGGCAAGATTCGCGTTCCACGGAATAT 3120
3121 AGAGGCCAACTTTCTCAATAGGTCTTGCAGAAACGAGAGCAGACTACACAGAGGCAAGTCT 3180
3121 AGAGGCCAACTTTCTCAATAGGTCTTGCAGAAACGAGAGCAGACTACACAGAGGCAAGTCT 3180
3181 CAACTTGCAGAGCTCTCGGTTAGTTAGCTTCAATGGAAATTTCTGACGTTATCTATAGAGA 3240
3181 CAACTTGCAGAGCTCTCGGTTAGTTAGCTTCAATGGAAATTTCTGACGTTATCTATAGAGA 3240
3241 GATCAATGCTCTCTTAACGTTATCTGCAATTCGCAATTCCTCTGCGGAAAGGAGCTT 3300
3241 GATCAATGCTCTCTTAACGTTATCTGCAATTCGCAATTCCTCTGCGGAAAGGAGCTT 3300
3301 CTAACACAGGTGTCTTCAAAGCGACTCCATCAAACTTGGCAGTTAGTTCTAAAAGGGCTT 3360
3301 CTAACACAGGTGTCTTCAAAGCGACTCCATCAAACTTGGCAGTTAGTTCTAAAAGGGCTT 3360
3361 TGTCAACATTTGACGAAACATTTGTCGCAATTTGGTTTGAATAATTCCTAATCTGTTCGG 3420
3361 TGTCAACATTTGACGAAACATTTGTCGCAATTTGGTTTGAATAATTCCTAATCTGTTCGG 3420
3421 TTTTCTGATAGCAGCAGCAAGGGCATCTTCAATTTCTTGTGAGGAGGCTTAGAAACGT 3480
3421 TTTTCTGATAGCAGCAGCAAGGGCATCTTCAATTTCTTGTGAGGAGGCTTAGAAACGT 3480
3481 CAATTTGCACAATTCOAATACGACCTTCAGAGGAGCTTCTTTAGTTGGATCTTCTT 3540
3481 CAATTTGCACAATTCOAATACGACCTTCAGAGGAGCTTCTTTAGTTGGATCTTCTT 3540
3541 TAGTTGTTTCTTGGTGTATCTTGGCTTGGCATCTCTTTCTCTTCTGAGTACCTTTAGGG 3600
3541 TAGTTGTTTCTTGGTGTATCTTGGCTTGGCATCTCTTTCTCTTCTGAGTACCTTTAGGG 3600
3601 ACTTCATATCCAGGTTTCTCTCCACTCGTCAACGCTCAACCGTACTTGGCACAATCTAA 3660
3601 ACTTCATATCCAGGTTTCTCTCCACTCGTCAACGCTCAACCGTACTTGGCACAATCTAA 3660
3661 CTAATCGAAATAAATAAATAGTCAGCACTTCCCGAGCTATATCTTCTTGGATTTAGCTT 3720
3661 CTAATCGAAATAAATAAATAGTCAGCACTTCCCGAGCTATATCTTCTTGGATTTAGCTT 3720
3721 CTGCAAGTTCTACAGCTTCTCTCCCTAATTTAGCGTTCAACAAAACCTTCTGCTCAAAATA 3780
3721 CTGCAAGTTCTACAGCTTCTCTCCCTAATTTAGCGTTCAACAAAACCTTCTGCTCAAAATA 3780
3781 ACCGTTTGGTATAAGAACCTTCTGAGCAATTCGCTTAGCATCCCAAGGTTGGCTTCCA 3840
3781 ACCGTTTGGTATAAGAACCTTCTGAGCAATTCGCTTAGCATCCCAAGGTTGGCTTCCA 3840
3841 TGGCTCTAAGACCCCTTGTGATTCGCAACAGAGAGTGGGTTCCAAAGTGCAGAAACCAA 3900
3841 TGGCTCTAAGACCCCTTGTGATTCGCAACAGAGAGTGGGTTCCAAAGTGCAGAAACCAA 3900
3901 CACCTGTTTGTTCACCAACCAATTTCAAGCAGCTCCATCAAACTCCAATTCGATACCCA 3960
3901 CACCTGTTTGTTCACCAACCAATTTCAAGCAGCTCCATCAAACTCCAATTCGATACCCA 3960
3961 GCAACTTTTGGTTCGTCAGATGTAGCACCTTTATACCAAAACCGTGACGACGAGATT 4020
3961 GCAACTTTTGGTTCGTCAGATGTAGCACCTTTATACCAAAACCGTGACGACGAGATT 4020
4021 GGTAGACTCCAGTTGTGCTTATAGCCTTCGGAATAGACTTTTGGACGAGTACACCA 4080
4021 GGTAGACTCCAGTTGTGCTTATAGCCTTCGGAATAGACTTTTGGACGAGTACACCA 4080
4081 GGGCCCAACAGTAATTTAGAAAGTGCAGCCACAAAGTAGTGAATAGACCATCGGGCGGT 4140
4081 GGGCCCAACAGTAATTTAGAAAGTGCAGCCACAAAGTAGTGAATAGACCATCGGGCGGT 4140

4141 CAGTAGTCAAGAGCGCAACAAAATTTCACTGACAGGAACTTTTGTACATCTTCAGAAA 4200
 Db CAGTAGTCAAGAGCGCAACAAAATTTCACTGACAGGAACTTTTGTACATCTTCAGAAA 4200
 4201 GTTCGATTCAGTAGTCAATTCGCGAGCATCAATATATGGGATATACAGAGCAACAG 4260
 Db GTTCGATTCAGTAGTCAATTCGCGAGCATCAATATATGGGATATACAGAGCAACAG 4260
 4261 TGGAGTCAATCTACCAACTTTTGGCGTCTCAGAAAAGCATAAAGTTCTACTACCGC 4320
 Db TGGAGTCAATCTACCAACTTTTGGCGTCTCAGAAAAGCATAAAGTTCTACTACCGC 4320
 4321 CATTAGTGAACCTTTCAATTCGCCAGTGGAGAAAGGCAAGCGATCTAGCAT 4380
 Db CATTAGTGAACCTTTCAATTCGCCAGTGGAGAAAGGCAAGCGATCTAGCAT 4380
 4381 TAGCGGCAAGGATGCAACTTTTCAACAGGCTCTATAGATAACCTAGCGCTGGGA 4440
 Db TAGCGGCAAGGATGCAACTTTTCAACAGGCTCTATAGATAACCTAGCGCTGGGA 4440
 4441 TCATCCCTTTGGACAACCTTTCTGCCAAATCTAGGTCGAAATCACTTCAATGATACCAT 4500
 Db TCATCCCTTTGGACAACCTTTCTGCCAAATCTAGGTCGAAATCACTTCAATGATACCAT 4500
 4501 TATTGTACAACTTACGCAAGTTGCGATCAGCTCCTCAAAATGGTCTCTGTAAACGATG 4560
 Db TATTGTACAACTTACGCAAGTTGCGATCAGCTCCTCAAAATGGTCTCTGTAAACGATG 4560
 4561 ACTCAACTTGCACATTAACCTTGAAGCTCAGTCGATGAACTTGATCAGGTTGTGCA 4620
 Db ACTCAACTTGCACATTAACCTTGAAGCTCAGTCGATGAACTTGATCAGGTTGTGCA 4620
 4621 GCTGGTCAGAGCATAGGGAACACGGCTTTTCTACCAACTCAAGGAATATCAAACT 4680
 Db GCTGGTCAGAGCATAGGGAACACGGCTTTTCTACCAACTCAAGGAATATCAAACT 4680
 4681 CTGCAACACTTGGCGATGAGCAAGGAAATGTCATCTTGAAGTCGACAGTGAG 4740
 Db CTGCAACACTTGGCGATGAGCAAGGAAATGTCATCTTGAAGTCGACAGTGAG 4740
 4741 TGTAGTCTTGAGAAATCTGAAGCGGATTTTATTATCAGTAGTCACTCAGAGAG 4800
 Db TGTAGTCTTGAGAAATCTGAAGCGGATTTTATTATCAGTAGTCACTCAGAGAG 4800
 4801 TCCTCTACGCGGACGATCGTGGCGACCTGACAGTCCGATCACCGCGCCACAGGTG 4860
 Db TCCTCTACGCGGACGATCGTGGCGACCTGACAGTCCGATCACCGCGCCACAGGTG 4860
 4861 CGGTTGCTGGCGCTTATATCCCGACATCACCGATGGGGAAGTCGGGCTCGCCACTTCG 4920
 Db CGGTTGCTGGCGCTTATATCCCGACATCACCGATGGGGAAGTCGGGCTCGCCACTTCG 4920
 4921 GGTCTACAGCGCTTTGTTTCGGCTGGGTATGGTGGCAGCCCGTGGCGGGGACTGT 4980
 Db GGTCTACAGCGCTTTGTTTCGGCTGGGTATGGTGGCAGCCCGTGGCGGGGACTGT 4980
 4981 TGGGGCCATCTCCTTGATGACCACTTCTTGGCGGGCGGTGCTCAACGGGCTCAACC 5040
 Db TGGGGCCATCTCCTTGATGACCACTTCTTGGCGGGCGGTGCTCAACGGGCTCAACC 5040
 5041 TACTACTGGCTGCTTCTTATGAGAGTGCATTAAGGAGAGCGTTCAGTATCTATGA 5100
 Db TACTACTGGCTGCTTCTTATGAGAGTGCATTAAGGAGAGCGTTCAGTATCTATGA 5100
 5101 TTGGAAGTATGGGAATGATACCCGATTTCTTCAAGTGTCTTCAAGTGTCTTCAAGT 5160
 Db TTGGAAGTATGGGAATGATACCCGATTTCTTCAAGTGTCTTCAAGTGTCTTCAAGT 5160
 5161 TATGCCCACTAAAGCAACCGAGAGGAGATTTCTAGGTAATTTCTCTGACTTTTGGT 5220
 Db TATGCCCACTAAAGCAACCGAGAGGAGATTTCTAGGTAATTTCTCTGACTTTTGGT 5220

5221 CATCAGTAGACTGAACTGTGAGACTATCTCGGTTATGACAGCAGAAAATCTCTCTCTGG 5280
 Db CATCAGTAGACTGAACTGTGAGACTATCTCGGTTATGACAGCAGAAAATCTCTCTCTGG 5280
 5281 AGACAGTAAATGAAGTCCCAACCAATAAAGAAATCTTGTATCAGGAAACAACTTCTTGT 5340
 Db AGACAGTAAATGAAGTCCCAACCAATAAAGAAATCTTGTATCAGGAAACAACTTCTTGT 5340
 5341 TTCGAACTTTTTCGGTCTTGAATATAAATCTAGTGGATATCTCGGTTAGGAATG 5400
 Db TTCGAACTTTTTCGGTCTTGAATATAAATCTAGTGGATATCTCGGTTAGGAATG 5400
 5401 GAGCGGCAAAATGCTTACCTTCTGGAACCTTCAAGAGGTATGTAGGTTTGTAGATACTGA 5460
 Db GAGCGGCAAAATGCTTACCTTCTGGAACCTTCAAGAGGTATGTAGGTTTGTAGATACTGA 5460
 5461 TGCCAACTTTCAGTGACAAAGTTCGTATTTCTCAACCAATTCGGATCCAGAGAAATCA 5520
 Db TGCCAACTTTCAGTGACAAAGTTCGTATTTCTCAACCAATTCGGATCCAGAGAAATCA 5520
 5521 AAGTTGTTTGTCTACTATTTGATCCAGCCAGTGGGCTTGAACCTGACAATAGTGTCT 5580
 Db AAGTTGTTTGTCTACTATTTGATCCAGCCAGTGGGCTTGAACCTGACAATAGTGTCT 5580
 5581 CGTGTGTTGAGGTCACTTTCGTATGAATTAATCTAGTCTTGTATCTAAATCTTGAACG 5640
 Db CGTGTGTTGAGGTCACTTTCGTATGAATTAATCTAGTCTTGTATCTAAATCTTGAACG 5640
 5641 AGCCAAAGGATTAATACCCAAATCTAAATCTTTTAAACCTTAAAGGACAAAGTATG 5700
 Db AGCCAAAGGATTAATACCCAAATCTAAATCTTTTAAACCTTAAAGGACAAAGTATG 5700
 5701 TCTGCTCTATTAACCCCAATCAGCTCGTAGTCTGATCTCTCACTCACTTGAAGGCGAC 5760
 Db TCTGCTCTATTAACCCCAATCAGCTCGTAGTCTGATCTCTCACTCACTTGAAGGCGAC 5760
 5761 TATCTGTTTGTAGAGAAATTTGCGAGATCGATTCGAGAAAGGTAGCTGATTTTA 5820
 Db TATCTGTTTGTAGAGAAATTTGCGAGATCGATTCGAGAAAGGTAGCTGATTTTA 5820
 5821 AACGTAAATTTATCTCAAGATCTCTGCTCGCGCTTTCGGTGTATGACGGTGAACCT 5880
 Db AACGTAAATTTATCTCAAGATCTCTGCTCGCGCTTTCGGTGTATGACGGTGAACCT 5880
 5881 CTGACATGACGCTCCCGAGACGGTTCACAGCTTGTCTGTAAGCGGATCCCGGAGCAG 5940
 Db CTGACATGACGCTCCCGAGACGGTTCACAGCTTGTCTGTAAGCGGATCCCGGAGCAG 5940
 5941 ACAAGCCGTGAGGCGGCTCAGCGGCTTTCGGGCTTTCGGGCTGATGACCGCATGACCCA 6000
 Db ACAAGCCGTGAGGCGGCTCAGCGGCTTTCGGGCTTTCGGGCTGATGACCGCATGACCCA 6000
 6001 GTACGTAGCGATAGCGGAGTGTATCTGCTTAACTATGCGGATCAGAGCAGATTGTA 6060
 Db GTACGTAGCGATAGCGGAGTGTATCTGCTTAACTATGCGGATCAGAGCAGATTGTA 6060
 6061 CTGAGAGTGCACCATATCGGCTGTAATACCGACAGATCGGTAAGGAGAAATACCGC 6120
 Db CTGAGAGTGCACCATATCGGCTGTAATACCGACAGATCGGTAAGGAGAAATACCGC 6120
 6121 ATCAGCGCTCTTTCGGCTTCTCTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6180
 Db ATCAGCGCTCTTTCGGCTTCTCTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6180
 6181 CGAGCGGTATCAGCTCACTCAAGCGGTATACCGGTTATCCACAGATCAGGGGATAC 6240
 Db CGAGCGGTATCAGCTCACTCAAGCGGTATACCGGTTATCCACAGATCAGGGGATAC 6240
 6241 GCAGGAAAGAAATCTGTAGCAAAAGCCAGCAAAAGCCAGGAAACCGTAAAGGCGCG 6300
 Db GCAGGAAAGAAATCTGTAGCAAAAGCCAGCAAAAGCCAGGAAACCGTAAAGGCGCG 6300
 6301 TTGCTGGCGCTTTTTCATAGGCTCGCCCCCTGACGAGCATCACAATAATCAGCGCTCA 6360

6301 TTGCTGGCGTTTTCATAGGCTCGGCCCTCGACGAGCATCAAAAATCGACGCTCA 6360
6361 AGTCAGAGGTGGCGAAACCCGACAGACTATAAGATACCGGCGTTTCCCGCTGGAGC 6420
6361 AGTCAGAGGTGGCGAAACCCGACAGACTATAAGATACCGGCGTTTCCCGCTGGAGC 6420
6421 TCCCTCGTGGCTCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCGCTTCTC 6480
6421 TCCCTCGTGGCTCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCGCTTCTC 6480
6481 CCGTCGGGAGCGTGGCGCTTCTCATAGCTACCGCTAGGTATCTCAGTTCGGTGTAG 6540
6481 CCGTCGGGAGCGTGGCGCTTCTCATAGCTACCGCTAGGTATCTCAGTTCGGTGTAG 6540
6541 GTGCTTCCGCTCCAGCTGGCTGTGTGCAAGAACCCCGCTTCAGCCGCAACCGCTCGGCC 6600
6541 GTGCTTCCGCTCCAGCTGGCTGTGTGCAAGAACCCCGCTTCAGCCGCAACCGCTCGGCC 6600
6601 TTATCCGGTAAGTATCTGTTAGTCCAAACCGGTAAGACAGCTTATCGCCACTGGCA 6660
6601 TTATCCGGTAAGTATCTGTTAGTCCAAACCGGTAAGACAGCTTATCGCCACTGGCA 6660
6661 GCAGCACTGGTAACAGGATAGCAGAGCGAGTATGTAGCGGTGTACAGAGTCTTGTG 6720
6661 GCAGCACTGGTAACAGGATAGCAGAGCGAGTATGTAGCGGTGTACAGAGTCTTGTG 6720
6721 AAGTGGTGGCTTAACCGCTACACTAGAGAGGAGTATTTGGTATCTCGGCTCTGCTG 6780
6721 AAGTGGTGGCTTAACCGCTACACTAGAGAGGAGTATTTGGTATCTCGGCTCTGCTG 6780
6781 AAGCAGATTACCTTCGGAAAGAGTGTGTAGTCTTGTATCGGCAACCAACACCGCT 6840
6781 AAGCAGATTACCTTCGGAAAGAGTGTGTAGTCTTGTATCGGCAACCAACACCGCT 6840
6841 GGTAGCGGTGGTGTGTTGTTGCAAGCAGCAGATTACCGCAGAAAAAAGGATCTCAA 6900
6841 GGTAGCGGTGGTGTGTTGTTGTTGCAAGCAGCAGATTACCGCAGAAAAAAGGATCTCAA 6900
6901 GAAGATCTTTGATCTTTTACGGGCTCAGCGCTAGTGAAGAGAACTACGTTAA 6960
6901 GAAGATCTTTGATCTTTTACGGGCTCAGCGCTAGTGAAGAGAACTACGTTAA 6960
6961 GGGATTTGGTTCATGAGATTAACAAAAAGGATCTTACCTAGATCTTTTAAATTAATA 7020
6961 GGGATTTGGTTCATGAGATTAACAAAAAGGATCTTACCTAGATCTTTTAAATTAATA 7020
7021 TGAAGTTTAAATCAATCTAAAGTATATAGTAAGTCTGGTCTGACAGTTACCAATGC 7080
7021 TGAAGTTTAAATCAATCTAAAGTATATAGTAAGTCTGGTCTGACAGTTACCAATGC 7080
7081 TTAATCAGTGGGACCTATCTCAGCGATCTGTATTTGGTTCATCATAGTTGCTGTA 7140
7081 TTAATCAGTGGGACCTATCTCAGCGATCTGTATTTGGTTCATCATAGTTGCTGTA 7140
7141 CTCCCGCTCGTGTAGATAACTACGATCGGAGGCTTACCATCTGGCCCCAGTGTGCA 7200
7141 CTCCCGCTCGTGTAGATAACTACGATCGGAGGCTTACCATCTGGCCCCAGTGTGCA 7200
7201 ATGATACCGGAGACCCAGCTCAGCGCTCCAGATTTATCAGCAATAAACCCAGCGCC 7260
7201 ATGATACCGGAGACCCAGCTCAGCGCTCCAGATTTATCAGCAATAAACCCAGCGCC 7260
7261 GGAAGGGCCGAGCGCAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTTAAAT 7320
7261 GGAAGGGCCGAGCGCAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTTAAAT 7320
7321 TGTTCGGGAGCTAGAGTAAAGTATGCGGAGTAAATAGTTTGGCAAGCTTTGTTGCC 7380
7321 TGTTCGGGAGCTAGAGTAAAGTATGCGGAGTAAATAGTTTGGCAAGCTTTGTTGCC 7380
7381 ATTGCTCAGGATCGTGGTGTACGCTCGCTTGTGTATGGTTCATTACAGCTCGGT 7440

7381 ATTGCTCAGGATCGTGGTGTCAAGCTCGTCTGTTGGTATGCTTCAATTCAGCTCGGT 7440
7441 TCCCAAGATCAAGGGGAGTTATCATGATCCCCCATGTTGCGAAAAAGCGGTAGCTCC 7500
7441 TCCCAAGATCAAGGGGAGTTATCATGATCCCCCATGTTGCGAAAAAGCGGTAGCTCC 7500
7501 TTCCGCTCTCCGATCGTTGTGCAAGTAAGTTTGGCGCGAGTGTATCACTCATGTTATG 7560
7501 TTCCGCTCTCCGATCGTTGTGCAAGTAAGTTTGGCGCGAGTGTATCACTCATGTTATG 7560
7561 GCAGCACTGATTAATTTCTTACTGTATGCCATCCGTAAGATGCTTTTCTGTGACTGTT 7620
7561 GCAGCACTGATTAATTTCTTACTGTATGCCATCCGTAAGATGCTTTTCTGTGACTGTT 7620
7621 GAGTACTCAACCAAGTCAATCTGAGATAGTGTATGCGCGACCGAGTGTCTTCCCGCG 7680
7621 GAGTACTCAACCAAGTCAATCTGAGATAGTGTATGCGCGACCGAGTGTCTTCCCGCG 7680
7681 GCCTCAACACGGGATTAATACCGCGCACATPAGCAGAACTTTAAAGTGTCTCATCTGGA 7740
7741 AAACGTTCTTCCGGGGGAAACCTCTCAAGATCTTACCGCTGTGAGATCCAGTTCGATG 7800
7741 AAACGTTCTTCCGGGGGAAACCTCTCAAGATCTTACCGCTGTGAGATCCAGTTCGATG 7800
7801 TAAACCACTCGTGCACCCAACTCTGAGATCTTACGATCTTTTACCTTACCAGCGTTCTGG 7860
7801 TAAACCACTCGTGCACCCAACTCTGAGATCTTACGATCTTTTACCTTACCAGCGTTCTGG 7860
7861 TGAGCAAAACAGGAGGCAAAATGCGCAAAAGGGAATAGCGGACGACGGAATGT 7920
7861 TGAGCAAAACAGGAGGCAAAATGCGCAAAAGGGAATAGCGGACGACGGAATGT 7920
7921 TGAATACTCATATCTCTTCTCTTCAATATTTAAGCAATTTATCAGGGTTATTTCTCTC 7980
7921 TGAATACTCATATCTCTTCTCTTCAATATTTAAGCAATTTATCAGGGTTATTTCTCTC 7980
7981 ATGAGCGGATACATATTTGAATGTTATGAAAAAATAACAAATAGGGGTTCCGCGCAC 8040
7981 ATGAGCGGATACATATTTGAATGTTATGAAAAAATAACAAATAGGGGTTCCGCGCAC 8040
8041 TTTCCCGGAAAAAGTGCACCTGACGCTAAGAAACCATTTATATCATGACATTAACCTAT 8100
8041 TTTCCCGGAAAAAGTGCACCTGACGCTAAGAAACCATTTATATCATGACATTAACCTAT 8100
8101 AAAAAATAGGCTATACAGAGGCGCTTCTTCAAGAAATTAATCTCATGTTTGCACG 8160
8101 AAAAAATAGGCTATACAGAGGCGCTTCTTCAAGAAATTAATCTCATGTTTGCACG 8160
8161 TTATCATGATAGCTGACTCATGTTGTTGTAATAGAGCGAGATCGGGAACACTG 8220
8161 TTATCATGATAGCTGACTCATGTTGTTGTAATAGAGCGAGATCGGGAACACTG 8220
8221 AAAAAATACAGTTATTATTTCG 8241
8221 AAAAAATACAGTTATTATTTCG 8241

RESULT 2
AAV28852
ID AAV28852 standard; cDNA to mRNA; 8598 BP.
XX
AC AAV28852;
XX
DT 04-AUG-1998 (first entry)
XX
DE pPic9/ELF25L cDNA construct containing the elafin 25L gene.
XX
KW Elafin; elafin 25L gene; Pichia pastoris; expression vector;
XX alcohol oxidase 1; controlling region; promoter; ds.
OS Synthetic.

OS Pichia pastoris.

XX Key Location/Qualifiers

PH CDS 949..1377

FT /*tag= a

FT sig_peptide 949..1203

FT /*tag= b

FT mat_peptide 1204..1374

FT /*tag= c

XX JP10127292-A.

PN

XX 19-MAY-1998.

PD

XX 31-OCT-1996; 96JP-00304233.

XX

XX 31-OCT-1996; 96JP-00304233.

XX (TSUR) TSUMURA & CO.

PA

XX WPI; 1998-340657/30.

DR

XX P-PSDB; AAM57237.

DR

XX New elafin expression vector - used to transform microbial host(s) for

PT production of elafin commercially.

PT

XX Example 1; Page 11-15; 21pp; Japanese.

PS

XX The present sequence represents a pPIC9/ELF25L cDNA construct containing

CC the elafin 25L gene, used in an example of the present invention. The

CC present invention describes: (i) an elafin (EL) expression vector

CC comprising a gene encoding EL and a gene encoding alcohol oxidase 1

CC controlling region (AOCR) which controls the expression of EL encoding

CC gene; (2) a microbial host transformed with the vector; (3) DNA fragment

CC for recombinant transformation comprising: (a) a first region homologous

CC to the genomic DNA of the host; (b) a promoter region of (AOCR) gene; (c)

CC a gene encoding signal peptide; (d) a linker designed to secrete a

CC matured elafin; (e) a gene encoding EL; (f) selection marker gene; (g)

CC stop codon, and (h) a second region homologous to a part of the genomic

CC DNA of the host, and (4) a microbial host transformed with the DNA

CC fragment of (3). The microbial hosts can be used to prepare EL

CC commercially. The production of EL by the microbial hosts can be carried

CC on a large scale

XX

SQ Sequence 8598 BP; 2306 A; 2025 C; 1920 G; 2347 T; 0 U; 0 Other;

Query Match 91.7%; Score 7555.6; DB 2; Length 8598;

Best Local Similarity 94.0%; Pred. No. 0;

Mismatches 8098; Conservative 0; Mismatches 129; Indels 385; Gaps 8;

QY 1 AGATCTACATCCAAAGACGAAAGTTCGAATGAACCTTTTGGCATCCGACATCCACAG 60

DB 1 AGATCTACATCCAAAGACGAAAGTTCGAATGAACCTTTTGGCATCCGACATCCACAG 60

QY 61 GTCCATTCTCACATAAGTGCACAAACGCAACAGGAGGGATACACTAGCAGACACCGT 120

DB 61 GTCCATTCTCACATAAGTGCACAAACGCAACAGGAGGGATACACTAGCAGACACCGT 120

QY 121 TCGCAACGCGAGACCTCCATCTCTCTCTCAACACCCACCTTTTGGCATCCGACATCCACAG 180

DB 121 TCGCAACGCGAGACCTCCATCTCTCTCTCAACACCCACCTTTTGGCATCCGACATCCACAG 180

QY 181 AGCCAGATTATTGGGCTTGAATGGAGCTCGCTCAATTCCTTCTATTAGGCTACTA 240

DB 181 AGCCAGATTATTGGGCTTGAATGGAGCTCGCTCAATTCCTTCTATTAGGCTACTA 240

QY 241 ACACCATGACTTTATTAGCTGCTCTATCTCTGGCCCTCGCGAGGTTCAATGTTCTTTA 300

DB 241 ACACCATGACTTTATTAGCTGCTCTATCTCTGGCCCTCGCGAGGTTCAATGTTCTTTA 300

QY 301 TTTCGGAATGCAACAGCTCCGCAATTACACCGAACATCACTCCAGATGAGGCTTTCTG 360

DB 301 TTTCGGAATGCAACAGCTCCGCAATTACACCGAACATCACTCCAGATGAGGCTTTCTG 360

QY 361 AGTGTGGGGTCAAAATAGTTTTCATGTTCCCAATGGCCCAAACTGACAGTTTAAACGCT 420

DB 361 AGTGTGGGGTCAAAATAGTTTTCATGTTCCCAATGGCCCAAACTGACAGTTTAAACGCT 420

QY 421 GTCTTTGGAACCTTAATGACAAAGCGTGTCTCATCCAAGATGAATAGTTTGGTTCTG 480

DB 421 GTCTTTGGAACCTTAATGACAAAGCGTGTCTCATCCAAGATGAATAGTTTGGTTCTG 480

QY 481 TTGGAATGCTTAACCGCCAGTTGGTCAAAAGAACTTCCAAAGTCCCATACCGTTTCT 540

DB 481 TTGGAATGCTTAACCGCCAGTTGGTCAAAAGAACTTCCAAAGTCCCATACCGTTTCT 540

QY 541 CTGTGTTGTTGATTGACGAATGCTCAAAAATAATCTCATTAATGCTTAGCGAGTCT 600

DB 541 CTGTGTTGTTGATTGACGAATGCTCAAAAATAATCTCATTAATGCTTAGCGAGTCT 600

QY 601 CTCTATCGCTTCTGAACCCCGGTGCACCTGTGCGGAAACGCAATGGGAAACACCCGCT 660

DB 601 CTCTATCGCTTCTGAACCCCGGTGCACCTGTGCGGAAACGCAATGGGAAACACCCGCT 660

QY 661 TTTTGGATGATTATGCAATGCTTCCACATGTTATGCTTCCAAAGATTTCTGGTGGGAATCT 720

DB 661 TTTTGGATGATTATGCAATGCTTCCACATGTTATGCTTCCAAAGATTTCTGGTGGGAATCT 720

QY 721 GCTGATAGCTTAACGTTTCATGATCAAAATTTAACTGTTCTAAACCCCTACTTGACAGCAAT 780

DB 721 GCTGATAGCTTAACGTTTCATGATCAAAATTTAACTGTTCTAAACCCCTACTTGACAGCAAT 780

QY 781 ATATAACAGAGGAAGCTGCCCTGTCTTAAACCTTTTATCATCATTTAGCTT 840

DB 781 ATATAACAGAGGAAGCTGCCCTGTCTTAAACCTTTTATCATCATTTAGCTT 840

QY 841 ACTTTCATAATGGGACTGGTCCCAATGACAGCTTTTGAATTTAACACTTTTAAACGA 900

DB 841 ACTTTCATAATGGGACTGGTCCCAATGACAGCTTTTGAATTTAACACTTTTAAACGA 900

QY 901 CAACTTGAGAGATCAAAAACCACTAATTAATTCGAAGGATCCAAACGATGAGATTCT 960

DB 901 CAACTTGAGAGATCAAAAACCACTAATTAATTCGAAGGATCCAAACGATGAGATTCT 960

QY 961 TCAATTTTACTGAGTTTATTCGCGATCTCTCGCATTTAGCTGCTCCAGTCAACACT 1020

DB 961 TCAATTTTACTGAGTTTATTCGCGATCTCTCGCATTTAGCTGCTCCAGTCAACACT 1020

QY 1021 ACAACGAGATGAACCGGACAAATTCGGGTGAAGCTGTCAATCGTTTACTCAGATTTA 1080

DB 1021 ACAACGAGATGAACCGGACAAATTCGGGTGAAGCTGTCAATCGTTTACTCAGATTTA 1080

QY 1081 GAAGGGATTTCCGATGTTGTTTCCCATTTTCCACAGCACAATAACGGTTATTG 1140

DB 1081 GAAGGGATTTCCGATGTTGTTTCCCATTTTCCACAGCACAATAACGGTTATTG 1140

QY 1141 TTTTATAATACTACTATTGCCAGCATTTGCTGTCTTAAAGAAAGAGGGGTATCTCTCGAGAA 1200

DB 1141 TTTTATAATACTACTATTGCCAGCATTTGCTGTCTTAAAGAAAGAGGGGTATCTCTCGAGAA 1200

QY 1201 AGAG---AGGCTGAAGCCCGGAGCCCGGAGCCCGGAGCCCGGAGGAAACATGATC 1257

DB 1201 AGAGCTCAGAAACGAGTTAAGGGTCCGGTGTGACCAAAACCGGGCTCTTGGCCGATTTC 1260

QY 1258 ATGCCCCCGGAGAGGATATAAATTGGGTTCCTCCGGTGTCCCGCCAGCAGTGCACG 1317

DB 1261 CTGATCGCTGCGCTTGTCTGAACCCGCGAACCGTTGTCTGAAGACACTGACTGCGCG 1320

QY 1318 GAG----- 1320

DB 1321 GGTAATCAAAAAATGCTGCGAAGGTTCTTGGGTATGGCATGCTTCTCCGCGAGTAGTGA 1380

QY 1321 ----- 1320

DB 1381 AGCTTCTCGAGGCCAGCTTGGCTGCGAGGTCGGAAGAACCCATCTGCCTAATGAGTAA 1440

QY 1321 ----- 1320
Db 1441 TATCTGATAAGAACGCTTTTATTTAGCTTTTAACTACATAGACGAATAAATATACGA 1500
QY 1321 ----- 1320
Db 1501 TCGGGATACATAGCATATTCAGCTCTTTTACCTCTCAATTTGCTCAAGATTATATCGT 1560
QY 1321 ----- 1320
Db 1561 GACCAATTTTAGATCATTTTCGGTCAGACAGCCACATTTTCGCGTAATTTTAAAGTTCAA 1620
QY 1321 ----- 1320
Db 1621 TTAATGTATGCGTGACACACTATATATAGCTCAACTGTGTAATAAGGAAGTAGTGATAAT 1680
QY 1355 TCCCGTGTGCTTCCACCCCATGGCCATCGAGAACACTCAAGA-----AGAAGAAATGTC 1409
Db 1681 AACGGTTTCAGGTTCCTCACATGCTTATTTCAAAATCTAAGATGTTTAAACATTTTGG 1740
QY 1410 CTTCTAACTAGTGGGT-----AGAA 1430
Db 1741 ATACTAACGACGATAGTTCTCTCGCATCTGCGACCGTTGATATTTAGTAAACTACAAGAA 1800
QY 1431 TTCCTTAGGGCGCGCAATTAATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTG 1490
Db 1801 TTCCTTAGGGCGCGCGCAATTAATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTG 1860
QY 1491 GGCACCTTAGGAGAACCGGCTTGTCTAGATTTCTAATCAAGAGGATGTCAGAAATGCCATT 1550
Db 1861 GGCACCTTAGGAGAACCGGCTTGTCTAGATTTCTAATCAAGAGGATGTCAGAAATGCCATT 1920
QY 1551 TGCCTGAGAGATGAGGCTTCATTTTGATGACTTTTATTTGTAACCTATATAGTATAG 1610
Db 1921 TGCCTGAGAGATGAGGCTTCATTTTGATGACTTTTATTTGTAACCTATATAGTATAG 1980
QY 1611 GATTTTTTTTGTCATTTTGTCTCTGACGAGCTTGCCTCGATCAGCCTATCTCGC 1670
Db 1981 GATTTTTTTTGTCATTTTGTCTCTGACGAGCTTGCCTCGATCAGCCTATCTCGC 2040
QY 1671 AGCTGATGAATATCTTTGTGGTAGGGTTTGGAAATCAATTCGAGTTTGATGTTTTCTT 1730
Db 2041 AGCTGATGAATATCTTTGTGGTAGGGTTTGGAAATCAATTCGAGTTTGATGTTTTCTT 2100
QY 1731 GGTATTTCCCACTCTCTTCAGAGTACAGAGATTAAGTGAGAGATTCGTTTGTGGCAAGC 1790
Db 2101 GGTATTTCCCACTCTCTTCAGAGTACAGAGATTAAGTGAGAGATTCGTTTGTGGCAAGC 2160
QY 1791 TTATCGATAAGCTTTTAAATCGGTAGTTTATCACAGTTAAATTTGCTAACGCACTCAGGCAC 1850
Db 2161 TTATCGATAAGCTTTTAAATCGGTAGTTTATCACAGTTAAATTTGCTAACGCACTCAGGCAC 2220
QY 1851 CGGTATGAATCTAACAAATGGCTTCATCTCGGACCGTCAACCGTCAACCGTCAACCGT 1910
Db 2221 CGGTATGAATCTAACAAATGGCTTCATCTCGGACCGTCAACCGTCAACCGTCAACCGT 2280
QY 1911 TAGGCATAGGCTTGGTTATGCGGTACTGCGGGGCTCTTGGGGATATCGTCCATTCGG 1970
Db 2281 TAGGCATAGGCTTGGTTATGCGGTACTGCGGGGCTCTTGGGGATATCGTCCATTCGG 2340
QY 1971 ACAGCATCGCCAGTCACTATGCGGTGCTGTAGCGCTATATCGGTGATGCAATTTCTAT 2030
Db 2341 ACAGCATCGCCAGTCACTATGCGGTGCTGTAGCGCTATATCGGTGATGCAATTTCTAT 2400
QY 2031 GGGCACCGGTTCTCGGACGACTGTCGACCGCTTTGGCGCGCGCCGCTGCTGCTCGCTT 2090
Db 2401 GGGCACCGGTTCTCGGACGACTGTCGACCGCTTTGGCGCGCGCCGCTGCTGCTCGCTT 2460
QY 2091 CGCTACTTTGGAGCCACTATGCACTAGCGATCATCGGACCAACACCGCTCTCTGAGTCT 2150
Db 2461 CGCTACTTTGGAGCCACTATGCACTAGCGATCATCGGACCAACACCGCTCTCTGAGTCT 2520
QY 2151 ATCGAATCTAATGTAAAGTTAAATCTCTAATATTAATTAAGTCCCGCTTTCTCCATA 2210

Db 2521 ATCGAATCTAATGTAAAGTTAAATCTCTAATATTAATTAAGTCCCGCTTTCTCCATA 2580
QY 2211 CGAACCTTAAACAGCATTTGGGTGAGCATCTAGACCTTCAACAGCAGCCAGATCCATCACT 2270
Db 2581 CGAACCTTAAACAGCATTTGGGTGAGCATCTAGACCTTCAACAGCAGCCAGATCCATCACT 2640
QY 2271 GCTTGCCCAATATGTTTTCAGTCCCTCAGGAGTTACGTTCTGTGAAGTGAAGAACTTCTGG 2330
Db 2641 GCTTGCCCAATATGTTTTCAGTCCCTCAGGAGTTACGTTCTGTGAAGTGAAGAACTTCTGG 2700
QY 2331 AAGTTTGCAGTGTAACTCGGTGATTTACGCGGCATATCCGTAAGTTCGCAAAAGTGTGG 2390
Db 2701 AAGTTTGCAGTGTAACTCGGTGATTTACGCGGCATATCCGTAAGTTCGCAAAAGTGTGG 2760
QY 2391 TTGCTACCGGAGGAGTAATCTCCAACTCTCTGGAGAGTAGGCACCAACAAACACAGAT 2450
Db 2761 TTGCTACCGGAGGAGTAATCTCCAACTCTCTGGAGAGTAGGCACCAACAAACACAGAT 2820
QY 2451 CAGCGTGTGTACTTGTATCAATAGAAAGCAATTCGATTTGCAAGGATCAAGTCT 2510
Db 2821 CAGCGTGTGTACTTGTATCAATAGAAAGCAATTCGATTTGCAAGGATCAAGTCT 2880
QY 2511 TCAGGAGCGTACTGATTTGGACATTTCCAAAGCCTGCTCGTAGGTTGCAACCGATAGGGTT 2570
Db 2881 TCAGGAGCGTACTGATTTGGACATTTCCAAAGCCTGCTCGTAGGTTGCAACCGATAGGGTT 2940
QY 2571 GTAGAGTGTCAATACATTTGCGTCAATTTCAACCTTTGGCAATCTGCAAGTGTGGTTG 2630
Db 2941 GTAGAGTGTCAATACATTTGCGTCAATTTCAACCTTTGGCAATCTGCAAGTGTGGTTG 3000
QY 2631 TGAACAGCATCTTCAATCTGCGCAAGCTCTTGTCTGTCATATCGACAGCAACAACATC 2690
Db 3001 TGAACAGCATCTTCAATCTGCGCAAGCTCTTGTCTGTCATATCGACAGCAACAACATC 3060
QY 2691 ACCTGGGAATCAATACCATTTGAGCTTGAGACAGAAAGTCTGAGGCAACGAAATCTGGA 2750
Db 3061 ACCTGGGAATCAATACCATTTGAGCTTGAGACAGAAAGTCTGAGGCAACGAAATCTGGA 3119
QY 2751 TCAGCGTATTTATCAGCAATACTAGAACTTCAGAGGGCCAGCAGGCAATGTCATACTA 2810
Db 3120 TCAGCGTATTTATCAGCAATACTAGAACTTCAGAGGGCCAGCAGGCAATGTCATACTA 3179
QY 2811 CACAGGCGTATGTCATTTGAACCATCATCTTGGCAGCAGTAACGAACTGTTTCTCT 2870
Db 3180 CACAGGCGTATGTCATTTGAACCATCATCTTGGCAGCAGTAACGAACTGTTTCTCT 3239
QY 2871 GGAACCAATATTTTGTGTCACATTTAGGAACAGTTTCTGTTCCGTAAGCCATAGCAGTCT 2930
Db 3240 GGAACCAATATTTTGTGTCACATTTAGGAACAGTTTCTGTTCCGTAAGCCATAGCAGTCT 3299
QY 2931 GCCTGGCGGCTCTCTGTCAGCATACACTTAGCAACCACTTGTGGCAACGTAAGTG 2990
Db 3300 GCCTGGCGGCTCTCTGTCAGCATACACTTAGCAACCACTTGTGGCAACGTAAGTG 3359
QY 2991 ACTTCTGGGTAAGGTAACATCTTCTTAGGTGGAGATGCAAAAACAAATTTCTTTGCAA 3050
Db 3360 ACTTCTGGGTAAGGTAACATCTTCTTAGGTGGAGATGCAAAAACAAATTTCTTTGCAA 3419
QY 3051 CCAGCACTTTGCGAGGAACACCCAGCATCAGGGAAGTGGAGGCAAGATTCGGGTTCCA 3110
Db 3420 CCAGCACTTTGCGAGGAACACCCAGCATCAGGGAAGTGGAGGCAAGATTCGGGTTCCA 3479
QY 3111 CCAGGAATATAGAGGCCAACTTTCTCAATAGGTCTTTCGCAAAACGAGAGCAGACTACACCA 3170
Db 3480 CCAGGAATATAGAGGCCAACTTTCTCAATAGGTCTTTCGCAAAACGAGAGCAGACTACACCA 3539
QY 3171 GGGCAAGTCTCAACTTTGCAACGCTCTCGTTAGTTCGTTAGCTTCATGGAAATTTCTGAGCTTA 3230
Db 3540 GGGCAAGTCTCAACTTTGCAACGCTCTCGTTAGTTCGTTAGCTTCATGGAAATTTCTGAGCTTA 3599
QY 3231 TCTATAGAGAGATCAATGGCTCTCTTAAAGTCTTCTGCAATTTGCAATAGTTCTCTCTGGG 3290

Db	3600	TCTATAGAGAGATCAATGGGCTCTTTAAAGTTATCTGGCAATTCGATAAGTTCTCTGGG	3659	QY	4370	GATACTAGCATTAGCGGCAAGAGATGCAATTTTATCAACGAGGTCCTATAGATAACCCCT	4429
QY	3291	AAAGGCGCTTCTAACACAGGTGCTTCAAGAGCACTCCATCAAACTTGGCAGTTAGTTCT	3350	Db	4739	GATACTAGCATTAGCGGCAAGAGATGCAATTTTATCAACGAGGTCCTATAGATAACCCCT	4798
Db	3660	AAAGGAGCTTCTAACACAGGTGCTTCAAGAGCACTCCATCAAACTTGGCAGTTAGTTCT	3719	QY	4430	AGCGCTGGGATCATCTTTGACAACTCTTTGCGCAAACTCTAGTCCCAAAATCACTTC	4489
QY	3351	AAAAGGGCTTTGTCCACATTTGACGAACTTGTGCAAACTTGGTGTGACTAATTCATA	3410	Db	4799	AGCGCTGGGATCATCTTTGACAACTCTTTGCGCAAACTCTAGTCCCAAAATCACTTC	4858
Db	3720	AAAAGGGCTTTGTCCACATTTGACGAACTTGTGCAAACTTGGTGTGACTAATTCATA	3779	QY	4490	ATTGATACCATTTATGTAACAACTTTGAGCAAGTTGTGATCAGCTCTCAAAATTTGGTCTC	4549
QY	3411	ATCTGTTCCGTTTCTGGATAGGACGAGAGGCACTTCAATTTCTTGTGAGGAGGCC	3470	Db	4859	ATTGATACCATTTATGTAACAACTTTGAGCAAGTTGTGATCAGCTCTCAAAATTTGGTCTC	4918
Db	3780	ATCTGTTCCGTTTCTGGATAGGACGAGAGGCACTTCAATTTCTTGTGAGGAGGCC	3839	QY	4550	TGTAACGAGTACACTCAACTTGCACATTAACCTTGAAGCTCAGTTCGATTTGAGTGAACCTT	4609
QY	3471	TTAGAAACGCTCAATTTTGCAAACTTCAATACGACCTTTCAGAGGGAATCTTTTGTGTTG	3530	Db	4919	TGTAACGAGTACACTCAACTTGCACATTAACCTTGAAGCTCAGTTCGATTTGAGTGAACCTT	4978
Db	3840	TTAGAAACGCTCAATTTTGCAAACTTCAATACGACCTTTCAGAGGGAATCTTTTGTGTTG	3899	QY	4610	CAGGTTGTCAGCTGGTCAGCAGCATAGGGAACAACGCGCTTTTCTACCAAACTCAAGGA	4669
QY	3531	GATTCCTTTTGTGTTTCTGGTGTATCTTGGCTTGGCATCTCTTCTCTCTAGTG	3590	Db	4979	CAGGTTGTCAGCTGGTCAGCAGCATAGGGAACAACGCGCTTTTCTACCAAACTCAAGGA	5038
Db	3900	GATTCCTTTTGTGTTTCTGGTGTATCTTGGCTTGGCATCTCTTCTCTCTAGTG	3959	QY	4670	ATTATCAAACTCTGCAACACTTGGTATGAGTAGCAAGGAATGTCTACTTTGAAGT	4729
QY	3591	ACCTTTAGGGAATTCATATCCAGGTTTCTCTCCACTCTGTCACAGTTCACACGTTACTTG	3650	Db	5039	ATTATCAAACTCTGCAACACTTGGTATGAGTAGCAAGGAATGTCTACTTTGAAGT	5098
Db	3960	ACCTTTAGGGAATTCATATCCAGGTTTCTCTCCACTCTGTCACAGTTCACACGTTACTTG	4019	QY	4730	CGGACAGTGAAGTGTAGTCTTCAAGAAATTTCTGAAGCCGTTATTTTATTATCAGTGAGT	4789
QY	3651	GCACTCTAACTAATGCAAAATAAATAAGTACGACATTCCTCCAGGCTATATCTTCTCTG	3710	Db	5099	CGGACAGTGAAGTGTAGTCTTCAAGAAATTTCTGAAGCCGTTATTTTATTATCAGTGAGT	5158
Db	4020	GCACTCTAACTAATGCAAAATAAATAAGTACGACATTCCTCCAGGCTATATCTTCTCTG	4079	QY	4790	TCATCAGAGATTCCTTACGCGGACGATCGTGGCGGACCTGCGAGTTCGCGATCAACGG	4849
QY	3711	GATTTAGTCTTCAAGTTTCAATCAGCTTCTCTCCCTAATTTTGTAGGTTTC-AACAAACTTC	3769	Db	5159	TCATCAGAGATTCCTTACGCGGACGATCGTGGCGGACCTGCGAGTTCGCGATCAACGG	5206
Db	4080	GATTTAGTCTTCAAGTTTCAATCAGCTTCTCTCCCTAATTTTGTAGGTTTC-AACAAACTTC	4139	QY	4850	CGCACAGTGGGTTGCTGGCGCTATATCGCGACATCACCGATGGGGAAGATCGGGC	4909
QY	3770	GTCTCTCAAAATACCGTTTGTGATTAAGAACTTTCTGGAGCAATGCTTTACGATCCCAAA	3829	Db	5207	CGCACAGTGGGTTGCTGGCGCTATATCGCGACATCACCGATGGGGAAGATCGGGC	5266
Db	4140	GTCTCTCAAAATACCGTTTGTGATTAAGAACTTTCTGGAGCAATGCTTTACGATCCCAAA	4199	QY	4910	TCGCCACTTCGCGCTCATAGAGCGCTTGTTCGCGGTGGGTATGTGGCAGGCCCGCTGGC	4969
QY	3830	GGTGGCTTCAATGCTTCAAGACCTTTGATTTGGCCCAAAACAGAGTGGTTCAGAGT	3889	Db	5267	TCGCCACTTCGCGCTCATAGAGCGCTTGTTCGCGGTGGGTATGTGGCAGGCCCGCTGGC	5326
Db	4200	GGT-GCTTTCATGGCTTAAGACCTTTGATTTGGCCCAAAACAGAGTGGTTCAGAGT	4258	QY	4970	CGGGGAGCTGTTGGGCGCATCTCTTCGATGACCATTTCTTCGCGCGCGGTGCTCAA	5029
QY	3890	ACAGAAACCAACCTGTTTCTCAACCAAAATTTCAAGCAGTCTCCATCACAATCCAA	3949	Db	5327	CGGGGAGCTGTTGGGCGCATCTCTTCGATGACCATTTCTTCGCGCGCGGTGCTCAA	5386
Db	4259	ACAGAAACCAACCTGTTTCTCAACCAAAATTTCAAGCAGTCTCCATCACAATCCAA	4318	QY	5030	CGGCCCTCAACCTTACTTACTTGGGCTGCTTCTTAATGAGGAGTTCGATTAAGGAGAGCGTCG	5089
QY	3950	TTCCGATACCCAGGCTTGTGATGTTGCTCGAGATGATAGCCTTTATACCAAAACCGTG	4009	Db	5387	CGGCCCTCAACCTTACTTACTTGGGCTGCTTCTTAATGAGGAGTTCGATTAAGGAGAGCGTCG	5446
Db	4319	TTCCGATACCCAGGCTTGTGATGTTGCTCGAGATGATAGCCTTTATACCAAAACCGTG	4378	QY	5090	AGTATCTATGATGGAAGTATGGGATGTTGATACCGCATTTCTTCAGTGTCTTGAGTGC	5149
QY	4010	ACGAGGATTTGGTAGCTCCAGTCTTGTCTTATAGCTTCCGGAATAGACTTTTGTGA	4069	Db	5447	AGTATCTATGATGGAAGTATGGGATGTTGATACCGCATTTCTTCAGTGTCTTGAGTGC	5506
Db	4379	ACGAGGATTTGGTAGCTCCAGTCTTGTCTTATAGCTTCCGGAATAGACTTTTGTGA	4438	QY	5150	TCCTATCAGATTTATCCCAACTTAAAGCAACCGGAGGAGATTTTCATGTTAAATTTCTC	5209
QY	4070	CGAGTACCAAGGCTTGTGATGTTGCTCGAGATGATAGCCTTTATACCAAAACCGTG	4129	Db	5507	TCCTATCAGATTTATCCCAACTTAAAGCAACCGGAGGAGATTTTCATGTTAAATTTCTC	5566
Db	4439	CGAGTACCAAGGCTTGTGATGTTGCTCGAGATGATAGCCTTTATACCAAAACCGTG	4498	QY	5210	TCATTTTGGTTCATCAGTAGACTCGAATGTGAGCTATCTCGGTTATGACGAGAAAT	5269
QY	4130	ATCGGGGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	4189	Db	5567	TCATTTTGGTTCATCAGTAGACTCGAATGTGAGCTATCTCGGTTATGACGAGAAAT	5626
Db	4499	ATCGGGGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	4558	QY	5270	GTCTCTTGGAGACAGTAAATGAAAGTCCCAATTAAGAAATTCCTTTGTTATCAGGAAC	5329
QY	4190	ATCTTTCAGAAAGTTCGATTTTCAATTTGCGAGCATCAATTAATGAGGATTTATACC	4249	Db	5627	GTCTCTTGGAGACAGTAAATGAAAGTCCCAATTAAGAAATTCCTTTGTTATCAGGAAC	5686
Db	4559	ATCTTTCAGAAAGTTCGATTTTCAATTTGCGAGCATCAATTAATGAGGATTTATACC	4618	QY	5330	AAACTTCTTTGTTTCGAACTTTTTCGGTGCCTTGAACCTATAAAATGAGTGGATATGTC	5389
QY	4250	AGAAGCAAGTGGAGTCAATCTACCACTTTGCGGTCTCAGAAAGATTAACAGT	4309	Db	5687	AAACTTCTTTGTTTCGAACTTTTTCGGTGCCTTGAACCTATAAAATGAGTGGATATGTC	5746
Db	4619	AGAAGCAAGTGGAGTCAATCTACCACTTTGCGGTCTCAGAAAGATTAACAGT	4678	QY	5390	GGTAGGATGAGCGGCAATGCTTACCTTCTGACCTTCAAGAGGTATGAGGGTTT	5449
QY	4310	TCTACTACCGCATTTAGTGAACCTTTTCAATTTGCGGATGAGGAGGAGGAGGAGGAGG	4369	Db	5747	GGTAGGATGAGCGGCAATGCTTACCTTCTGACCTTCAAGAGGTATGAGGGTTT	5806
Db	4679	TCTACTACCGCATTTAGTGAACCTTTTCAATTTGCGGATGAGGAGGAGGAGGAGGAGG	4738				

QY 5450 GTAGATCTGATGCAACTTCAGTGACAAAGTTCGTTCTATTTCTGTTCAAAACCAATCCGAATC 5509
Db 5807 GTAGATCTGATGCCAACTTCAGTGACAAAGTTCGTTCTATTTCTGTTCAAAACCAATCCGAATC 5866
QY 5510 CAGAGAAATCAAGTTGTTGTTCTACTATTCAGTCCAGCCAGTGGGTTCTGAACTGAC 5569
Db 5867 CAGAGAAATCAAGTTGTTGTTCTACTATTCAGTCCAGCCAGTGGGTTCTGAACTGAC 5926
QY 5570 AATAGTGTGCTGCTGTTTGTAGGTCATCTTTGTATGAATAAATCTAGTCTTTGATCTAAA 5629
Db 5927 AATAGTGTGCTGCTGTTTGTAGGTCATCTTTGTATGAATAAATCTAGTCTTTGATCTAAA 5986
QY 5630 TAACTCTCAGCAGCCAGCGGATTAATACCAAACTCTTAACTCTTAAACCTTAAACCTTAAAC 5689
Db 5987 TAACTCTCAGCAGCCAGCGGATTAATACCAAACTCTTAAACCTCTTAAACCTTAAACCTTAAAC 6046
QY 5690 GGACAGATATGTCCTGTTTAAACCCCAAACTCAGCTCGTAGTCTGATCCTCATCAAC 5749
Db 6047 GGACAGATATGTCCTGTTTAAACCCCAAACTCAGCTCGTAGTCTGATCCTCATCAAC 6106
QY 5750 TTGAGGGGCACTATCTTGTAGAGAAATTTGCGGAGATCGGATATCGAGAAAGGTA 5809
Db 6107 TTGAGGGGCACTATCTTGTAGAGAAATTTGCGGAGATCGGATATCGAGAAAGGTA 6166
QY 5810 CGCTGATTTAAACGTGAAATTTATCTCAAGATCTCTGCTCGCGGTTTCGGTGATGAC 5869
Db 6167 CGCTGATTTAAACGTGAAATTTATCTCAAGATCTCTGCTCGCGGTTTCGGTGATGAC 6226
QY 5870 GGTGAAACCTCTGACATGACATGCTCCGAGACGGTCAAGCTTGTCTTAAACCGGAT 5929
Db 6227 GGTGAAACCTCTGACATGACATGCTCCGAGACGGTCAAGCTTGTCTTAAACCGGAT 6286
QY 5930 GCGGGAGCAGACAGCCGTCAGGCGGTCAGCGGTCAGCGGTCAGCGGTCAGCGGTCAGCGGTCAG 5989
Db 6287 GCGGGAGCAGACAGCCGTCAGGCGGTCAGCGGTCAGCGGTCAGCGGTCAGCGGTCAGCGGTCAG 6346
QY 5990 GCATGACCCAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6049
Db 6347 GCATGACCCAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6406
QY 6050 AGCAGAT 6109
Db 6407 AGCAGAT 6466
QY 6110 GAAATACCGATCAGGCGCTCTCCGCTCTCTCCGCTCACTGATGCTGCTGCTGCTGCTGCTGCTG 6169
Db 6467 GAAATACCGATCAGGCGCTCTCCGCTCTCTCCGCTCACTGATGCTGCTGCTGCTGCTGCTGCTG 6526
QY 6170 TTCGCTGCGGAGCGGTATCAGTCTCACTCAAGGCGGTATATCGGTTATCCACAGAT 6229
Db 6527 TTCGCTGCGGAGCGGTATCAGTCTCACTCAAGGCGGTATATCGGTTATCCACAGAT 6586
QY 6230 CAGGGGATACCGAGGAGACATGTCAGCAAAAGGCGCAGAAAGGCGCAGGACCGTA 6289
Db 6587 CAGGGGATACCGAGGAGACATGTCAGCAAAAGGCGCAGAAAGGCGCAGGACCGTA 6646
QY 6290 AAAAGGCGCGTGTGCTGCGTTTTCATAGGCTCCGCCCCCTGACGAGCATCACAAA 6349
Db 6647 AAAAGGCGCGTGTGCTGCGTTTTCATAGGCTCCGCCCCCTGACGAGCATCACAAA 6706
QY 6350 ATCGACGCTCAAGTCAGAGTGGGAGAACCCAGACGACTATAAGATACAGGCGTTTC 6409
Db 6707 ATCGACGCTCAAGTCAGAGTGGGAGAACCCAGACGACTATAAGATACAGGCGTTTC 6766
QY 6410 CCCCTGGAAGCTCCCTCGTGGCTCTCTGTTTCCGACCTCGGCTTACCGGATACCTGT 6469
Db 6767 CCCCTGGAAGCTCCCTCGTGGCTCTCTGTTTCCGACCTCGGCTTACCGGATACCTGT 6826
QY 6470 CCGCCTTCTCCCTTGGGAGAGCGTGGCTTTCTCAATAGCTCAAGCTGTAGGTATCTCA 6529
Db 6827 CCGCCTTCTCCCTTGGGAGAGCGTGGCTTTCTCAATAGCTCAAGCTGTAGGTATCTCA 6886
QY 6530 GTTCGGTGTAGTGTGCTCCAGCTGGGCTGTGTGCAAGAACCCCGGTTCCAGCCG 6589

Db 6887 GTTCGGTGTAGTGTGCTCCAGCTGGGCTGTGTGCAAGAACCCCGGTTCCAGCCG 6946
QY 6590 ACCGCTGCGCTTATCCGGTAACTATCGTCTTTAGTCCAAACCCGGTAAAGACAGACTTAT 6649
Db 6947 ACCGCTGCGCTTATCCGGTAACTATCGTCTTTAGTCCAAACCCGGTAAAGACAGACTTAT 7006
QY 6650 CGCCACTGCGAGCAGCAGCTGCTGACAGGATTTAGCAGCAGGAGGTATGTAGCGGCTGCTA 6709
Db 7007 CGCCACTGCGAGCAGCAGCTGCTGACAGGATTTAGCAGCAGGAGGTATGTAGCGGCTGCTA 7066
QY 6710 CAGAGTCTTGAAGTGTGCTGCTTAACTACGGCTTACACTAGAAAGCAGATTTTGTGTATCT 6769
Db 7067 CAGAGTCTTGAAGTGTGCTGCTTAACTACGGCTTACACTAGAAAGCAGATTTTGTGTATCT 7126
QY 6770 GCCTCTGTGTAAGCAGTACCTTCGGAAGAGTGTGCTGCTTGTATCCGCAAC 6829
Db 7127 GCCTCTGTGTAAGCAGTACCTTCGGAAGAGTGTGCTGCTTGTATCCGCAAC 7186
QY 6830 AAACCCCGCTGTAGCGGTGTTTGTGTTTGAAGCAGCAGATTTAGCGCAGAAAAA 6889
Db 7187 AAACCCCGCTGTAGCGGTGTTTGTGTTTGAAGCAGCAGATTTAGCGCAGAAAAA 7246
QY 6890 AAGGATCTCAAGAGATCTTGTGATCTTTCTACGGGTCGACGCTCAGTGGAAAGAA 6949
Db 7247 AAGGATCTCAAGAGATCTTGTGATCTTTCTACGGGTCGACGCTCAGTGGAAAGAA 7306
QY 6950 ACTCAGTGAAGGATTTGCTCATGAGATTTCAAAAGATCTTCACTAGATCTTT 7009
Db 7307 ACTCAGTGAAGGATTTGCTCATGAGATTTCAAAAGATCTTCACTAGATCTTT 7366
QY 7010 TAAATTAATAAAGTGTGTTTAAATCAATCTAAAGTATATATGATGATGATGATGATGATGATGAT 7069
Db 7367 TAAATTAATAAAGTGTGTTTAAATCAATCTAAAGTATATATGATGATGATGATGATGATGATGAT 7426
QY 7070 GTTACCAATGCTTAACTCAGTGGGACCTTATCTCAGCGATCTGCTATTTCTGTTTCACTCA 7129
Db 7427 GTTACCAATGCTTAACTCAGTGGGACCTTATCTCAGCGATCTGCTATTTCTGTTTCACTCA 7486
QY 7130 TAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7189
Db 7487 TAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7546
QY 7190 CCAGTGTGCAATGATACCGGAGACCCAGCTCAGCGCTCCAGATTTATCAGCAATAA 7249
Db 7547 CCAGTGTGCAATGATACCGGAGACCCAGCTCAGCGCTCCAGATTTATCAGCAATAA 7606
QY 7250 ACCAGCAGCGGAGGCGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7309
Db 7607 ACCAGCAGCGGAGGCGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7666
QY 7310 AGTCTATTATTTGTCGGGAAAGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7369
Db 7667 AGTCTATTATTTGTCGGGAAAGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7726
QY 7370 AGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7429
Db 7727 AGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7786
QY 7430 TCAGTCTCGGCTTCCCAACGATCAAGCGAGTTCATGATGATGATGATGATGATGATGATGATGATGATGAT 7489
Db 7787 TCAGTCTCGGCTTCCCAACGATCAAGCGAGTTCATGATGATGATGATGATGATGATGATGATGATGAT 7846
QY 7490 CGGTAGCTCTCTCGGCTTCCGATCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 7549
Db 7847 CGGTAGCTCTCTCGGCTTCCGATCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 7906
QY 7550 TCATGTTTATGCGCAGCAGCTGATTAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7609
Db 7907 TCATGTTTATGCGCAGCAGCTGATTAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7966
QY 7610 CTGTGCTGTGATGATCTCAACCAAGTCAATCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 7669

7967	CTGTGACTGGTGTAGTACTCAACCAAGTCATTCTGAGAAATAGTGTATTCGGCGGACCGAGTT	8022
7670	GCTCTTCCCGGGCGTCAACACACGGGATAAATACCGCGGCCACATAGCAGAACTTTAAAAAGTGC	7729
8027	GCTCTTCCCGGGCGTCAAACACGGGATANTACCGCGGCCACATAGCAGAACTTTAAAAAGTGC	8086
7730	TCATCATTTGGAAAACGTTCTTTCCGGCGGAAAACTCTCAAGGATCTTACCGCTCTTGAGAT	7789
8087	TCATCATTTGGAAAACGTTCTTTCCGGCGGAAAACTCTCAAGGATCTTACCGCTCTTGAGAT	8146
7790	CCAGTTTCGATGTAAACCCATCTCGTGCACCCAACTGTATCTTCAGCATCTTTTACTTTTCAACCA	7849
8147	CCAGTTTCGATGTAAACCCATCTCGTGCACCCAACTGTATCTTCAGCATCTTTTACTTTTCAACCA	8206
7850	GCCTTTCTGGGTGAGCAAAAAACAGGAAAGCAAAATGCCGAAAAAGGAATAAGGGCGA	7909
8207	GCCTTTCTGGGTGAGCAAAAAACAGGAAAGCAAAATGCCGCAAAAAGGGAATAAGGGCGA	8266
7910	CACGGAATGTTGAAATACTCATATCTCTTCCTTTTTCAAATATATTGAAGCAATTTATCAGG	7969
8267	CACGGAATGTTGAAATACTCATATCTCTTCCTTTTTCAAATATATTGAAGCAATTTATCAGG	8326
7970	GTTATTGTCATCAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATATAGGGG	8029
8327	GTTATTGTCATCAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATATAGGGG	8386
8030	TTCCGGCGCACATTTCCCGGAAAAAGTGCACCTGACGCTTAAGAAAAACCATTTATCATGA	8089
8387	TTCCGGCGCACATTTCCCGGAAAAAGTGCACCTGACGCTTAAGAAAAACCATTTATCATGA	8446
8090	CATTAACTTATAAAAATATAGCGGTATCAGGAGGCCCTTTCGTCTTCAAGAAATTAATTCTCA	8149
8447	CATTAACTTATAAAAATATAGCGGTATCAGGAGGCCCTTTCGTCTTCAAGAAATTAATTCTCA	8506
8150	TGTTTCACAGCTTATCATCCATTAAGCTGACTCATGTGTGGTATTTGTGAAATAGACGCGAGAT	8209
8507	TGTTTGACAGCTTATCATCCATTAAGCTGACTCATGTGTGGTATTTGTGAAATAGACGCGAGAT	8566
8210	CGGGAACACTGAAAAATAACAGTTATTATTTCG	8241
8567	CGGGAACACTGAAAAATAACAGTTATTATTTCG	8598

FT	mutation	1237. .1239	
FT		/*tag= d	
FT		/note= "CTG to TTC alteration of codon usage"	
FT	mutation	1243. .1245	
FT		/*tag= e	
FT		/note= "GGC to GGT alteration of codon usage"	
FT	mutation	1246. .1248	
FT		/*tag= f	
FT		/note= "ACC to ACT alteration of codon usage"	
FT	mutation	1249. .1251	
FT		/*tag= g	
FT		/note= "ATC to ATT alteration of codon usage"	
FT	mutation	1252. .1254	
FT		/*tag= h	
FT		/note= "AAA to AAG alteration of codon usage"	
FT	mutation	1255. .1257	
FT		/*tag= i	
FT		/note= "GCC to GCT alteration of codon usage"	
FT	mutation	1258. .1260	
FT		/*tag= j	
FT		/note= "GGC to GGT alteration of codon usage"	
FT	mutation	1261. .1263	
FT		/*tag= k	
FT		/note= "GTA to GTT alteration of codon usage"	
FT	mutation	1264. .1266	
FT		/*tag= l	
FT		/note= "GAA to GAG alteration of codon usage"	

[illegible]

[illegible]

1141	QY	TTTTATAATACTACTATTGCCAGCATGTGCTAAAGAGAGAGGGGTATCTCTCGAGAA	1201
1141	DB	TTTTATAATACTACTATTGCCAGCATGTGCTAAAGAGAGAGGGGTATCTCTCGAGAA	1201
1201	QY	AGAGAGGCTCAAGCCAGGCCAGGCCCAGGCCAGCG--	1236
1201	DB	AGAGAGGCTCAAGGCTTACGTAGAAATTCGACGTACATTTGTAAGGTACTATTAAAGGCTGGT	1260
1237	QY	-----	1236
1261	DB	GTTGAGACTTCCCGCTCTGTATTTCCACCAGAACGGCCAAAGTTACTGAAGTTACAACCGCT	1320
1237	QY	-----	1236
1321	DB	ACCGGCATCGTTGATTTGGGTTTCGAAAATCGGCTTCAAAGGCCAAGAAGACCTCGGTAAAC	1380
1237	QY	-----	1236
1381	DB	GGCCTGAAAGCCATTTGGCAGGTTGAGCAAAAAGACATCTATCGCCGGTACTGACTCCGGT	1440
1237	QY	-----	1236
1441	DB	TGGGGCAACCGCCAAATCCTTCATCGCGTTGAAAGGGCGCTTCGGTAAATTCGCGCTCGGT	1500
1237	QY	-----GCCAGGAAGAAAACATGTATCATGGCCCCCGGGAGAGAGATAA-----	1279
1501	DB	CGTTTGAACAGCGTCTCTGAAAGACACGGCGGACATCAATCTCTTGGGATAGCAAAAGCGAC	1560
1280	QY	-----	1279
1561	DB	TATTTGGGTATAACCAAAATTCGCGAACCCGAGGCGACGCTCATTTTCGGTACGCTACGAT	1620
1280	QY	-----	1279
1621	DB	TCTCCCGAATTTGCCGCTCAGGGCGAGGTACAAATACGGGCTTAACGACATGCGAGC	1680
1280	QY	-----	1279
1681	DB	AGACATAACAGCGAATCTTACCAGCGCGCTTCAACTACAAAAACGGTGGCTTCTTCGTG	1740
1280	QY	-----	1279
1741	DB	CAATATGCGCGTGCTATAAAGACATCATCAAGTGCAAGAGGGCTTGAATATGTAGAAA	1800
1280	QY	-----	1279
1801	DB	TACCAGATTACCGTTTGGTTCAGCGGTTACGACAATGATGCCCTGTACGCTTCCGTAGCC	1860
1280	QY	-----	1279
1861	DB	GTACAGCAACAAGACGCGAAACTGACTGATGCTTCCAATTGCGACAACCTCTCAACCGAA	1920
1280	QY	-----ATTGTGGCTTCCCGGGTGTCAACGCCCGCAGC1309	1309
1921	DB	GTTCCGCTACCTTGGCATACCGCTTCGGCAAGTAACGCCCGGAGTTTCTTACGCCAC	1980
1310	QY	AGTGACCGAGAGAGGTGTGTTTTGATGACAGTGTCCG-----	1349
1981	DB	GGCTTCAAAAGTTTGGTTGTATGATGCAGACATAGGCAACGAATACGACCAAGTGGTGTGTC	2040
1350	QY	-----GGATTCGCGTGGTGTCTCCACCGCCATGCCCATCGA1385	1385
2041	DB	GGTGGGAATACGACTTCTCTCAAACGCGACTTCGCTTGGTTTCTCGCGGTTGGTTGCAA	2100
1386	QY	GAAACACTCAAGAAGAAGAAATGTCCCTTCTTAACTAGTGGCG-----TA1427	1427
2101	DB	GAAGGCAAGGCGAAACAAATTCGTAGCGACTGCGCGCGGTGTGTGTCTGCTGCACAAA	2160
1428	QY	GAATTCCTTAGGGCGCGCGGAATTAATTCGCTTAGACATGACTGTTCTCTCAATTCGAAG	1487
2161	DB	TTCTAACTTAGGGCGCGCGGAATTAATTCGCTTAGACATGACTGTTCTCTCAATTCGAAG	2220
1488	QY	TTGGGCACTTACGAGAGACCGGTCTTGTAGATTTCAATCAAGAGGATGTCAAGATGCC	1547

Qy 3708 TTGGATTTAGCTTTCGCAAGTTTCATCAGCTTCTCCCTAAATTTTAGCGTTCAACAAACT 3767
Db 4441 TTGGATTTAGCTTTCGCAAGTTTCATCAGCTTCTCCCTAAATTTTAGCGTTCAACAAACT 4500
Qy 3768 TCGTCGTCAAAATAACCGTTTGGTATAAGAACTTCTGGAGCATGCTCTTACGATCCAC 3827
Db 4501 TCGTCGTCAAAATAACCGTTTGGTATAAGAACTTCTGGAGCATGCTCTTACGATCCAC 4560
Qy 3828 AAGTGGCTTCCATGGCTCTAAGACCTTTGATTGGCCAAAACAGGAAGTCCGTTCCAAG 3887
Db 4561 AAGTGGCTTCCATGGCTCTAAGACCTTTGATTGGCCAAAACAGGAAGTCCGTTCCAAG 4620
Qy 3888 TGACAGAAACCAACACCTTTGTTTCAACCAAAATTTCAAGCATGCTCCATCAATCC 3947
Db 4621 TGACAGAAACCAACACCTTTGTTTCAACCAAAATTTCAAGCATGCTCCATCAATCC 4680
Qy 3948 AATTCGATACCCAGCAACTTTTGTAGTCTCCAGATGTAGCACCTTTATATACCAAAACCG 4007
Db 4681 AATTCGATACCCAGCAACTTTTGTAGTCTCCAGATGTAGCACCTTTATATACCAAAACCG 4740
Qy 4008 TGACGACGAGATTTGTAGACTCCAGTTTGTCTTTATAGCTCCGGAATAGACTTTTGT 4067
Db 4741 TGACGACGAGATTTGTAGACTCCAGTTTGTCTTTATAGCTCCGGAATAGACTTTTGT 4800
Qy 4068 GACGATACACAGCCCAACAGTAATTAGAAGTTCAGCCACCAAGTAGTAGATA 4127
Db 4801 GACGATACACAGCCCAACAGTAATTAGAAGTTCAGCCACCAAGTAGTAGATA 4860
Qy 4128 CCATCGGGGGCTCAGTAGTCAAGACGCCCAAAATTTCACTGACAGGCACTTTTGT 4187
Db 4861 CCATCGGGGGCTCAGTAGTCAAGACGCCCAAAATTTCACTGACAGGCACTTTTGT 4920
Qy 4188 ACATCTTCAGAAAGTTCTGTAATTCAGTAGTCAATTTGCCGAGCATCAATATGGGATATA 4247
Db 4921 ACATCTTCAGAAAGTTCTGTAATTCAGTAGTCAATTTGCCGAGCATCAATATGGGATATA 4980
Qy 4248 CCAGAAGCAACAGTGGAGTCACTACCACTTTGGGTCTCAGAAAAGCATAACA 4307
Db 4981 CCAGAAGCAACAGTGGAGTCACTACCACTTTGGGTCTCAGAAAAGCATAACA 5040
Qy 4308 GTTCTACTACCGCAATTAGTGAATTTTCAATCGCCCACTGGAGAGAAAAGGCACA 4367
Db 5041 GTTCTACTACCGCAATTAGTGAATTTTCAATCGCCCACTGGAGAGAAAAGGCACA 5100
Qy 4368 GCGATCTAGANTTAGCGGCAAGATGCACTTTATCAACGAGGTCCTATAGTAACC 4427
Db 5101 GCGATCTAGANTTAGCGGCAAGATGCACTTTATCAACGAGGTCCTATAGTAACC 5160
Qy 4428 CTAGCGCTGGGATCATCTTTGGACAACCTTTCTGCAAAATCTAGGTCCAAAATCACT 4487
Db 5161 CTAGCGCTGGGATCATCTTTGGACAACCTTTCTGCAAAATCTAGGTCCAAAATCACT 5220
Qy 4488 TCATTTGATACANTTATGTACAACTTGGCAAGTTGTGATCAGCTCCAAATGGTCC 4547
Db 5221 TCATTTGATACANTTATGTACAACTTGGCAAGTTGTGATCAGCTCCAAATGGTCC 5280
Qy 4548 TCTGTAAACGATGACTCAACTTGCACATTAATTCAGCTCAGTCGATGAGTGAACCTTG 4607
Db 5281 TCTGTAAACGATGACTCAACTTGCACATTAATTCAGCTCAGTCGATGAGTGAACCTTG 5340
Qy 4608 ATCAGGTTGTGAGTGTGAGCAGCAGTAGGAAACACCGGTTTTTCTTACCAAACTCAAG 4667
Db 5341 ATCAGGTTGTGAGTGTGAGCAGCAGTAGGAAACACCGGTTTTTCTTACCAAACTCAAG 5400
Qy 4668 GAATTTATCAACTCTGCAACCTTCGATCGAGTACGAGGAAATGTCTACTTTGAA 4727
Db 5401 GAATTTATCAACTCTGCAACCTTCGATCGAGTACGAGGAAATGTCTACTTTGAA 5460
Qy 4728 GTCGACAGTGTAGTGTGAGAAATTTCTGAAGCCGTATTTTATATCAGTGAGTC 4787
Db 5461 GTCGACAGTGTAGTGTGAGAAATTTCTGAAGCCGTATTTTATATCAGTGAGTC 5520

Qy 4788 AGTCATCAGAGATCTCTACGCCGACGATCGTGGCCGACCTGCAGGTTCGCATCAC 4847
Db 5521 AGTCATCAGAGATCTCTACGCCGACGATCGTGGCCGACCTGCAGGTTCGCATCAC 5580
Qy 4848 GGGCCACACAGTGGTGGTGGCGCTATATGCCGACATCACCGATGGGAGATCG 4907
Db 5581 GGGCCACACAGTGGTGGTGGCGCTATATGCCGACATCACCGATGGGAGATCG 5640
Qy 4908 GCTCGGCACCTTCGGGCTCATGAGCGCTTGTTCGGCGTGGTATGTTGCGAGCCCGTG 4967
Db 5641 GCTCGGCACCTTCGGGCTCATGAGCGCTTGTTCGGCGTGGTATGTTGCGAGCCCGTG 5700
Qy 4968 GCGGGGGACTGTTGGCGGCACTCTCTTCGATGCAACATTCCTCGCGGCGCGTGTCTC 5027
Db 5701 GCGGGGGACTGTTGGCGGCACTCTCTTCGATGCAACATTCCTCGCGGCGCGTGTCTC 5760
Qy 5028 AACGGCTCAACCTACTACTTGGTCTCTTAAATGACAGAGTCGATTAAGGAGAGCT 5087
Db 5761 AACGGCTCAACCTACTACTTGGTCTCTTAAATGACAGAGTCGATTAAGGAGAGCT 5820
Qy 5088 CGAGTATCTATGATTCGAAATGGAATGGTATACCCGCAATTCCTCAGTGTCTTGAGG 5147
Db 5821 CGAGTATCTATGATTCGAAATGGAATGGTATACCCGCAATTCCTCAGTGTCTTGAGG 5880
Qy 5148 TCTCCTATCAGATATGCCCACTRAAGCAACCGAGAGGAGATTCATGTAATTTTC 5207
Db 5881 TCTCCTATCAGATATGCCCACTRAAGCAACCGAGAGGAGATTCATGTAATTTTC 5940
Qy 5208 TCTGACTTTTGGTTCATCAGTAGACTCGAACTGTGAGACTATCTCGGTTATGACAGAGAA 5267
Db 5941 TCTGACTTTTGGTTCATCAGTAGACTCGAACTGTGAGACTATCTCGGTTATGACAGAGAA 6000
Qy 5268 ATGTCCTTCTTGAGACAGTAATGAATGCCCAACCAATTAAGAAATCTTGTATCAGGA 5327
Db 6001 ATGTCCTTCTTGAGACAGTAATGAATGCCCAACCAATTAAGAAATCTTGTATCAGGA 6060
Qy 5328 ACAAATCTTGTTCGAACTTTTCGGTGCCTTGAATCTATAAATGTAGAGTGGATATG 5387
Db 6061 ACAAATCTTGTTCGAACTTTTCGGTGCCTTGAATCTATAAATGTAGAGTGGATATG 6120
Qy 5388 TCGGTAGGAATCGAGCGGCAAACTCTTCTGACCTTCTGACCTTCAAGAGGTATGAGGT 5447
Db 6121 TCGGTAGGAATCGAGCGGCAAACTCTTCTGACCTTCTGACCTTCAAGAGGTATGAGGT 6180
Qy 5448 TTCTAGATCTGATGCACTTCAGTGACAAAGTCTTATTCGTTCAAAACCAATCCGAA 5507
Db 6181 TTGTAGATCTGATGCACTTCAGTGACAAAGTCTTATTCGTTCAAAACCAATCCGAA 6240
Qy 5508 TCCAGAGAAATCAAAGTTGTTGTCTACTATTTGATCCAGCCAGTCCGGTCTTGAAACTG 5567
Db 6241 TCCAGAGAAATCAAAGTTGTTGTCTACTATTTGATCCAGCCAGTCCGGTCTTGAAACTG 6300
Qy 5568 ACAAATAGTGTCTGTTTGGAGTCTATCTTGTATGATAAATCTAGTCTTTGATCTA 5627
Db 6301 ACAAATAGTGTCTGTTTGGAGTCTATCTTGTATGATAAATCTAGTCTTTGATCTA 6360
Qy 5628 AATAATCTTGCAGGCAAGGCGATATAACCAAACTCTTAAACTCTTTTAAACCTTAA 5687
Db 6361 AATAATCTTGCAGGCAAGGCGATATAACCAAACTCTTAAACTCTTTTAAACCTTAA 6420
Qy 5688 AAGGACAAATGATGTCCTGTATTAACCCCAATCAGCTCGTAGTCTGATCTCATCA 5747
Db 6421 AAGGACAAATGATGTCCTGTATTAACCCCAATCAGCTCGTAGTCTGATCTCATCA 6480
Qy 5748 ACTTGGGGGCACTATCTTGTTCAGAGAAATTTGGGAGATCGGATTCGAGAAAAGG 5807
Db 6481 ACTTGGGGGCACTATCTTGTTCAGAGAAATTTGGGAGATCGGATTCGAGAAAAGG 6540
Qy 5808 TAGCTGATTTTAAACGTAATTTATCTCAAGATCTCTGCTCGCGCTTTTCGGTATG 5867
Db 6541 TAGCTGATTTTAAACGTAATTTATCTCAAGATCTCTGCTCGCGCTTTTCGGTATG 6600
Qy 5868 ACGGTGAACCTCTGACACATGCAGCTCCCGGAGAGCGTCAAGCTTCTCTGTAAGCGG 5927

QY 8088 GACATTAACCTATAAAATAGCGGTATCAACGAGCCCTTTGCTTCAAGATTAATCT 8147
DB 8821 GACATTAACCTATAAAATAGCGGTATCAACGAGCCCTTTGCTTCAAGATTAATCT 8880
QY 8148 CATGTTTGACAGCTTATCATCGATAAGCTGACTCATGTTGTTGTAATGGAATAGACGAG 8207
DB 8891 CATGTTTGACAGCTTATCATCGATAAGCTGACTCATGTTGTTGTAATGGAATAGACGAG 8940
QY 8208 ATCGGGAACACTGAAAAATAACAGTTATTAATTCG 8241
DB 8941 ATCGGGAACACTGAAAAATAACAGTTATTAATTCG 8974
RESULT 4
AAF87135
ID AAF87135 standard; DNA; 9408 BP.
XX
AC AAF87135;
XX
DT 26-MAR-2002 (first entry)
XX
DE Lunasin coding sequence.
XX
KW Lunasin; cancer; human; colon; upper gastrointestinal tract; breast;
KW prostate; liver; kidney; therapy; soybean; ds.
XX
OS Glycine max.
XX
FN WO200172784-A2.
XX
PD 04-OCT-2001.
XX
PF 23-MAR-2001; 2001WO-US009453.
XX
PR 24-MAR-2000; 2000US-00534705.
XX
PA (FILI-) FILIGEN BIOSCIENCES INC.
XX
PI Galvez AF;
XX
DR WPI; 2001-648438/74.
XX
PT Treatment or prevention of cancer by administering a lunasin peptide
PT which binds to non-acetylated histones H3 and H4 and prevents their
PT acetylation in hypoacetylated repressed chromatin.
XX
PS Disclosure; Page 43-46; 49pp; English.
XX
CC This sequence encodes a lunasin peptide used in the method of the
CC invention. The method is for the treatment or prevention of cancer,
CC comprising administration of a lunasin peptide or its active fragment or
CC analog. The invention is used to treat or prevent cancer, particularly in
CC a human. The cancers that can be treated and prevented include those of
CC the colon, upper gastrointestinal tract, breast, prostate, liver, kidney
CC or any other internal organs or tissues
XX
SQ Sequence 9408 BP; 2533 A; 2233 C; 2117 G; 2525 T; 0 U; 0 Other;
Query Match 81.2%; Score 6687.6; DB 4; Length 9408;
Best Local Similarity 85.2%; Pred No. 0;
Matches 8083; Conservative 0; Mismatches 84; Indels 1315; Gaps 3;
QY 1 AGATCTAATCAACGAAGAGAAAGTTGAATGAACCTTTTGGCATCCGACATCCACAG 60
DB 1 AGATCTAATCAACGAAGAGAAAGTTGAATGAACCTTTTGGCATCCGACATCCACAG 60
QY 61 GTCCATTCTCACATAAGTCCAAACGACAGGGGGATACACTAGCAGCAGCCGT 120
DB 61 GTCCATTCTCACATAAGTCCAAACGACAGGGGGATACACTAGCAGCAGCCGT 120
QY 121 TGCAACGACGAGACCTCCACTCTCTCTCCCTCAACACCCACTTTTGCCATCGAAAAACC 180

DB 121 TGCAACGACGAGACCTCCACTCTCTCTCTCCCTCAACACCCACTTTTGCCATCGAAAAACC 180
QY 181 AGCCAGTTTATTTGGGCTTGATTGGAGCTCGCTCATTTCCAAATTCCTTCTATAGGCTACTA 240
DB 181 AGCCAGTTTATTTGGGCTTGATTGGAGCTCGCTCATTTCCAAATTCCTTCTATAGGCTACTA 240
QY 241 ACACCATGCTTTATTTAGGCTCTCTATCTGCGCCCTCCCTGGCGAGGTTTCATGTTGTTTA 300
DB 241 ACACCATGCTTTATTTAGGCTCTCTATCTGCGCCCTCCCTGGCGAGGTTTCATGTTGTTTA 300
QY 301 TTTCCGAATGCAACCAAGCTCCGATTACACCCGAACATCACTCCAGATGAGGGCTTTCTG 360
DB 301 TTTCCGAATGCAACCAAGCTCCGATTACACCCGAACATCACTCCAGATGAGGGCTTTCTG 360
QY 361 AGTGTGGGTCAAATAGTTTTCATGTTCCCAATATGGCCCAAACTGACAGTTTAAACGCT 420
DB 361 AGTGTGGGTCAAATAGTTTTCATGTTCCCAATATGGCCCAAACTGACAGTTTAAACGCT 420
QY 421 GTCTTGAAACCTAATATGACAAAAGCGTGATCTCATCCAAGATGAACCTAAGTTTGGTTGG 480
DB 421 GTCTTGAAACCTAATATGACAAAAGCGTGATCTCATCCAAGATGAACCTAAGTTTGGTTGG 480
QY 481 TTGAATGCTAACCGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCGCCATACCGTTTGT 540
DB 481 TTGAATGCTAACCGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCGCCATACCGTTTGT 540
QY 541 CTGTTTGGTATTGATTGACGAATGCTCAAAAATTAATCTCATTATGCTTACGCGAGTCT 600
DB 541 CTGTTTGGTATTGATTGACGAATGCTCAAAAATTAATCTCATTATGCTTACGCGAGTCT 600
QY 601 CTCTATCGCTTCTGAAACCCCGGTGCACTGTCGCCGAACGCAATGGGGAACACACCCGCT 660
DB 601 CTCTATCGCTTCTGAAACCCCGGTGCACTGTCGCCGAACGCAATGGGGAACACACCCGCT 660
QY 661 TTTTGGATGATTATGCAATTCCTCCACATTTGATGTTCCAGATTCCTGGTGGGAACT 720
DB 661 TTTTGGATGATTATGCAATTCCTCCACATTTGATGTTCCAGATTCCTGGTGGGAACT 720
QY 721 GCTGATAGCTTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTGACAGCAAT 780
DB 721 GCTGATAGCTTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTGACAGCAAT 780
QY 781 ATATAACAGAGAGAGAGCTGCCCTGTCTTAAACCTTTTATCATCATTTATTTAGCTT 840
DB 781 ATATAACAGAGAGAGAGCTGCCCTGTCTTAAACCTTTTATCATCATTTATTTAGCTT 840
QY 841 ACTTTCATTAATTGCGACTGTTTCCAAATTCGCAAGCTTTTGATTTTAAACGATTTTAAACGA 900
DB 841 ACTTTCATTAATTGCGACTGTTTCCAAATTCGCAAGCTTTTGATTTTAAACGATTTTAAACGA 900
QY 901 CAACTTGAGAAGATCAAAAAACAACTAATTATCGAAGGATCCAAACGATGAGATTTCT 960
DB 901 CAACTTGAGAAGATCAAAAAACAACTAATTATTCGAAGGATCCAAACGATGAGATTTCT 960
QY 961 TCAATTTTACTGCACTTTTATTCGAGATTCCTCGGATGAGTCTCCAGCTCAACACT 1020
DB 961 TCAATTTTACTGCACTTTTATTCGAGATTCCTCCGAGATTCCTCCGATGAGTCTCCAGCTCAACACT 1020
QY 1021 ACAACAGAAGATGAACCGCACAAATTCGCGTGAAGCTGTGTCATCGGTTTACTCAGATTTA 1080
DB 1021 ACAACAGAAGATGAACCGCACAAATTCGCGTGAAGCTGTGTCATCGGTTTACTCAGATTTA 1080
QY 1081 GAAGGGATTCGATGTTGCTGTTTGGCAATTTTCCACAGCAGCAATAAACCGGTTATTG 1140
DB 1081 GAAGGGATTCGATGTTGCTGTTTGGCAATTTTCCACAGCAGCAATAAACCGGTTATTG 1140
QY 1141 TTTATTAATACACTACTATTCGCGACTGTTGCTGCTAAAGAGAGGGGTATCTCTCGAGAAA 1200
DB 1141 TTTATTAATACACTACTATTCGCGACTGTTGCTGCTAAAGAGAGGGGTATCTCTCGAGAAA 1200
QY 1201 AGAGAGGCTGAAGCCCGAGCCCGAGCCCGAGCCCGAGCCCGAGAGAAACATGTTATCATG 1260
DB 1201 AGAGAGGCTGAAGCTTACGTAT---CCAAATGGCAGCACCAGATAGTGGCGCAAG 1257

QY	1261	GC	CCCCCGGAGAGGATAAATTGTGGCTTCCC	CGGTGTACCCGCCAGAGTGCACGGAG	1320	QY	2341	TG	TAACTCCGCTGTATTGACGGGCATATCCG	TACGTGTGCAAAAGTGTGTTGGTACCGG	2400
DB	1258	CAG	CTCCAGGGGGTGAACCTCAGCCCTCGGAG	AGCACATCATGAGAGATCCAAAGC	1317	DB	2267	TG	TAACTCCGCTGTATTGACGGGCATATCCG	TACGTGTGCAAAAGTGTGTTGGTACCGG	2326
QY	1321	AG	AGTGTGCTTTTGTGATGACAGTGTCCGGG	GATCCCGTGTCTCCACCCATGGCC	1380	QY	2401	AG	AGTAAATCTCCAACTCTCTCGAGAGTAGG	CACCAACAAACACAGATCCAGCGTGT	2460
DB	1318	CG	CGCGCATGACGATGATGATGAC-----	-----	1344	DB	2327	AG	AGTAAATCTCCAACTCTCTCGAGAGTAGG	CACCAACAAACACAGATCCAGCGTGT	2386
QY	1381	AT	CGAGAACACTCAAGAGAGAAATGCCCTT	CTAACTAGTGGCGTAGAATCCCTTAGGG	1440	QY	2461	GT	ACTTGATCAACATAGAAGACATTTCTCGAT	TTGCAAGATCAAGTGTTCAGGAGCGT	2520
DB	1345	-----	-----GACGACTAAGAAATCCCTTAGGG	-----	1366	DB	2387	GT	ACTTGATCAACATAGAAGACATTTCTCGAT	TTGCAAGATCAAGTGTTCAGGAGCGT	2446
QY	1441	CG	CGCGCGAATTAATTCGCCTTAGACATGACT	GTCTCTCAGTTCAAGTTGGGCACATTAGC	1500	QY	2521	ACT	GATTGGAATTTCCAAAGCCTGCTCGTAGT	TTGCAACCGATAGGTTGTAGAGTGTG	2580
DB	1367	CG	CGCGCGGAATTAATTCGCCTTAGACATGACT	GTCTCTCAGTTCAAGTTGGGCACATTAGC	1426	DB	2447	ACT	GATTGGAATTTCCAAAGCCTGCTCGTAGT	TTGCAACCGATAGGTTGTAGAGTGTG	2506
QY	1501	AGA	AGACCGGTCTTGCTAGATTCTTAATCAAG	AGGATGTCAAGATGCCATTTGCTCGAGAG	1560	QY	2581	CA	ATACACTTCGGTACAAATTTCAACCCCTT	GGCAACTGCACAGCTTGGTTGTGAACAGCAT	2640
DB	1427	AGA	AGACCGGTCTTGCTAGATTCTTAATCAAG	AGGATGTCAAGATGCCATTTGCTCGAGAG	1486	DB	2507	CA	ATACACTTCGGTACAAATTTCAACCCCTT	GGCAACTGCACAGCTTGGTTGTGAACAGCAT	2566
QY	1561	AT	GAGCGCTCAATTTTGTATCACTTTTATTT	TATTTGTAACCTATATAGTATAGGATTTT	1620	QY	2641	CT	TCAATTTCTGGCAAGCTCTTTGTCTGTCA	TATCGACAGCAACAGAAATCACCTGGGAAT	2700
DB	1487	AT	GAGCGCTCAATTTTGTATCACTTTTATTT	TATTTGTAACCTATATAGTATAGGATTTT	1546	DB	2567	CT	TCAATTTCTGGCAAGCTCTTTGTCTGTCA	TATCGACAGCAACAGAAATCACCTGGGAAT	2626
QY	1621	GT	CAATTTGTTTCTTCGACGAGCTTGCTCTG	ATCAGCCTATCTCGCAGCTGATGAA	1680	QY	2701	CA	ATACCAATTTGAGCTTGTGACAGAGAGGT	CTGAGGCAACGAAATCTGGATCAGCGTATT	2760
DB	1547	GT	CAATTTGTTTCTTCGACGAGCTTGCTCTG	ATCAGCCTATCTCGCAGCTGATGAA	1606	DB	2627	CA	ATACCAATTTGAGCTTGTGACAGAGAGGT	CTGAGGCAACGAAATCTGGATCAGCGTATT	2686
QY	1681	TAT	CTTGTGTAGGGTTGGGAAATCAATTCGAG	TTTGATGTTTTTCTTGTGATTTCCC	1740	QY	2761	TAT	CAGCAATAAATAGAACTTCAGAAAGCC	CCAGCAGCATGTCAATACTACACAGGCTG	2820
DB	1607	TAT	CTTGTGTAGGGTTGGGAAATCAATTCGAG	TTTGATGTTTTTCTTGTGATTTCCC	1666	DB	2687	TAT	CAGCAATAAATAGAACTTCAGAAAGCC	CCAGCAGCATGTCAATACTACACAGGCTG	2746
QY	1741	ACT	CCTCTTCAGAGTACAGAGATTAAGTGA	AGATTTCGTTGTGCAAGCTTATCGATAA	1800	QY	2821	AT	GTCTCAATTTGAGCAATCATCTTTGGCA	GAGTAAACGAACTGTTTCTGGACCAATA	2880
DB	1667	ACT	CCTCTTCAGAGTACAGAGATTAAGTGA	AGATTTCGTTGTGCAAGCTTATCGATAA	1726	DB	2747	AT	GTCTCAATTTGAGCAATCATCTTTGGCA	GAGTAAACGAACTGTTTCTGGACCAATA	2806
QY	1801	GCT	TTAATCGCGTATTTATCACAGTTAAAT	TGCTAAACGAGTACAGCAACGTTGATGAA	1860	QY	2881	TTTT	GTCTCACACTTAGAAACAGTTTCTGTT	CCGTAAGCCATAGCAGTACTCCCTGGGCGC	2940
DB	1727	GCT	TTAATCGCGTATTTATCACAGTTAAAT	TGCTAAACGAGTACAGCAACGTTGATGAA	1786	DB	2807	TTTT	GTCTCACACTTAGAAACAGTTTCTGTT	CCGTAAGCCATAGCAGTACTCCCTGGGCGC	2866
QY	1861	AT	CTAAACATCGCTCATCGTCCTCGCAGC	CGTCAACCTGATGCTGTAGGCAATAGG	1920	QY	2941	CT	CTCTGTAGCAGCATACACTTAGCACC	CAACCTTTGTGGGCAACGTAGATGCTTCTGGG	3000
DB	1787	AT	CTAAACATCGCTCATCGTCCTCGCAGC	CGTCAACCTGATGCTGTAGGCAATAGG	1846	DB	2867	CT	CTCTGTAGCAGCATACACTTAGCACC	CAACCTTTGTGGGCAACGTAGATGCTTCTGGG	2926
QY	1921	CT	TGTTATGCGGTACTGCGCGGCTCTTGT	CGGGATATCGTCCATCCGACAGCATCGC	1980	QY	3001	TA	AGGGTACCACTCTTCTTAGTGTGAGAT	GCAAAACAAATTTCTTGCAACACAGCACTT	3060
DB	1847	CT	TGTTATGCGGTACTGCGCGGCTCTTGT	CGGGATATCGTCCATCCGACAGCATCGC	1906	DB	2927	TA	AGGGTACCACTCTTCTTAGTGTGAGAT	GCAAAACAAATTTCTTGCAACACAGCACTT	2986
QY	1981	CAG	TCACTATGGGTGTGTAGCGCTATATGCG	TTGATGCAATTTCTATCGCACCCGT	2040	QY	3061	TG	CGAGGAACCCACGATCAGGGAAGTGG	GAAGGAGAAATTCGGTTCACACAGGAATAT	3120
DB	1907	CAG	TCACTATGGGTGTGTAGCGCTATATGCG	TTGATGCAATTTCTATCGCACCCGT	1966	DB	2987	TG	CGAGGAACCCACGATCAGGGAAGTGG	GAAGGAGAAATTCGGTTCACACAGGAATAT	3046
QY	2041	TCT	CGAGCATGTCCGACCGCTTTGGCGCG	CGCCAGTCTGCTCGCTTGGTACTTGG	2100	QY	3121	AG	AGCGCAACTTTCTCAATAGTCTTTGCA	AAACCGAGCAGACTACACAGCGCAAGTCT	3180
DB	1967	TCT	CGAGCATGTCCGACCGCTTTGGCGCG	CGCCAGTCTGCTCGCTTGGTACTTGG	2026	DB	3047	AG	AGCGCAACTTTCTCAATAGTCTTTGCA	AAACCGAGCAGACTACACAGCGCAAGTCT	3106
QY	2101	AG	CCACTATCGACTACGGATCATGGCGG	ACACACCGCTCTGTGGATCTATCGAATCTA	2160	QY	3181	CA	ACTTGAACCTGTCCGTTAGTGTGAGCT	TCATGGAATTTCTGACGTTATCTATAGAGA	3240
DB	2027	AG	CCACTATCGACTACGGATCATGGCGG	ACACACCGCTCTGTGGATCTATCGAATCTA	2086	DB	3107	CA	ACTTGAACCTGTCCGTTAGTGTGAGCT	TCATGGAATTTCTGACGTTATCTATAGAGA	3166
QY	2161	AT	GTAAAGTTAAATCTCTAAATTAATTA	ATTAAGTCCAGTTTCTCCATAGCACTTAA	2220	QY	3241	GA	TCAATGGCTCTCTTAACGTTATCTG	CAATTCGCAATTCCTCTGGGAAGGAGCTT	3300
DB	2087	AT	GTAAAGTTAAATCTCTAAATTAATTA	ATTAAGTCCAGTTTCTCCATAGCACTTAA	2146	DB	3167	GA	TCAATGGCTCTCTTAACGTTATCTG	CAATTCGCAATTCCTCTGGGAAGGAGCTT	3226
QY	2221	CAG	CAATTCGGGTGAGACTAGAACCTTCA	ACAGCAGCAGCATCCATCACTGCTTGGCCAA	2280	QY	3301	CT	AAACAGGTGCTCTTCAAAAGCGACTCC	ATCAAACTTGGCAGTTAGTTCTAAAAAGGGCTT	3360
DB	2147	CAG	CAATTCGGGTGAGACTAGAACCTTCA	ACAGCAGCAGCATCCATCACTGCTTGGCCAA	2206	DB	3227	CT	AAACAGGTGCTCTTCAAAAGCGACTCC	ATCAAACTTGGCAGTTAGTTCTAAAAAGGGCTT	3286
QY	2281	TAT	TTTTTCACTCCAGGATGATGCTCTG	TGAAGTATCACTCTGGAAGTTGACG	2340	QY	3361	TG	TCACTTTTGAACGAAATTTGTGCA	AAATTTGGTTTCAATTCATATCTGTTCCG	3420
DB	2207	TAT	TTTTTCACTCCAGGATGATGCTCTG	TGAAGTATCACTCTGGAAGTTGACG	2266	DB	3287	TG	TCACTTTTGAACGAAATTTGTGCA	AAATTTGGTTTCAATTCATATCTGTTCCG	3346
						DB	3421	TTTT	CTGTGATAGACGACGAAGGGCATCTT	CAATTTCTTGTGAGGAGGCTTTAGAAACGT	3480

Db	3347	TTTTCTGGATAGGACGAGGAAGGCATCTTCAA	TTCTTTGTGAGGAGGCCTTTAGAAACGT	3406
Qy	3481	CAATTTTGCACAAATCAATACGACCTTTCAGAAGGAC	TTCTTTTAGTGTGGATTTCTTCTT	3540
Db	3407	CAATTTTGCACAAATCAATACGACCTTTCAGAAGGAC	TTCTTTTAGTGTGGATTTCTTCTT	3466
Qy	3541	TAGGTTGTTCTTGGTGATCCTGGCTTGGCATCTCCCT	TTCTTCTTAGTGACCTTTTAGG	3600
Db	3467	TAGGTTGTTCTTGGTGATCCTGGCTTGGCATCTCCCT	TTCTTCTTAGTGACCTTTTAGG	3526
Qy	3601	ACTTCATATCCAGGTCTTCTTCCACCTCGTCCAAAGT	CACACCGCTACTTGGCACATCTAA	3660
Db	3527	ACTTCATATCCAGGTCTTCTTCCACCTCGTCCAAAGT	CACACCGCTACTTGGCACATCTAA	3586
Qy	3661	CTAATGCAAAATAAAATAAGTCAGCACATTC	CCAGGCTATATCTTCTTGGATTTAGCTT	3720
Db	3587	CTAATGCAAAATAAAATAAGTCAGCACATTC	CCAGGCTATATCTTCTTGGATTTAGCTT	3646
Qy	3721	CTGCAAGTTTCATCAGCTTCCTCCCTTAATTTTAG	CGTTTCAACAAACCTCGTCGTCAAATA	3780
Db	3647	CTGCAAGTTTCATCAGCTTCCTCCCTTAATTTTAG	CGTTTCAACAAACCTCGTCGTCAAATA	3706
Qy	3781	ACCGTTTGGTATAGAACCCTTTTGGAGCATTCCT	TACGATCCCAAGTGGCTTCCA	3840
Db	3707	ACCGTTTGGTATAGAACCCTTTTGGAGCATTCCT	TACGATCCCAAGTGGCTTCCA	3766
Qy	3841	TGGCTCTAAGACCTTTGATTGGCCAAAACAGGAAG	TGCGTTCCAAAGTGCAGACACAA	3900
Db	3767	TGGCTCTAAGACCTTTGATTGGCCAAAACAGGAAG	TGCGTTCCAAAGTGCAGACACAA	3826
Qy	3901	CACCTGTTTGTTCAACACACAAATTTCAAGCAGT	CTCCATCACAATCCAAATTCGATACCCA	3960
Db	3827	CACCTGTTTGTTCAACACACAAATTTCAAGCAGT	CTCCATCACAATCCAAATTCGATACCCA	3886
Qy	3961	GCAACTTTTGAGTTGCTCCAGATGTAGACCTTTT	ATACCAAAACCGTGACGACGAGATT	4020
Db	3887	GCAACTTTTGAGTTGCTCCAGATGTAGACCTTTT	ATACCAAAACCGTGACGACGAGATT	3946
Qy	4021	GGTAGACTCCAGTTTGTGTCCTTATAGCCTCCGG	AATAGACTTTTGGACGAGTACACCA	4080
Db	3947	GGTAGACTCCAGTTTGTGTCCTTATAGCCTCCGG	AATAGACTTTTGGACGAGTACACCA	4006
Qy	4081	GGCCCAACGAGTAAATTAGAAAGTTCAGCCACCA	AAAGTAGTGAATAGACCAATCGGGGCGGT	4140
Db	4007	GGCCCAACGAGTAAATTAGAAAGTTCAGCCACCA	AAAGTAGTGAATAGACCAATCGGGGCGGT	4066
Qy	4141	CAGTAGTCAAAAGACCCCAACAAATTTTCACTG	ACAGGGACCTTTTGGACATCTTCAGAA	4200
Db	4067	CAGTAGTCAAAAGACCCCAACAAATTTTCACTG	ACAGGGACCTTTTGGACATCTTCAGAA	4126
Qy	4201	GTTTCGTATTTCAGTAGTCAATTTGCCGAGCAT	CAATTAATGGGGATTATACCAAGACCAACAG	4260
Db	4127	GTTTCGTATTTCAGTAGTCAATTTGCCGAGCAT	CAATTAATGGGGATTATACCAAGACCAACAG	4186
Qy	4261	TGGNAGTCATCTACCAACTTTTGGCGTCTCAG	AAAAGCATAAACGTTCTACTACCGC	4320
Db	4187	TGGNAGTCATCTACCAACTTTTGGCGTCTCAG	AAAAGCATAAACGTTCTACTACCGC	4246
Qy	4321	CATTAGTGAACACTTTTCAAATTCGCCCAGTGG	AGAGAAAAAGGCACAGCGATACTAGCAT	4380
Db	4247	CATTAGTGAACACTTTTCAAATTCGCCCAGTGG	AGAGAAAAAGGCACAGCGATACTAGCAT	4306
Qy	4381	TAGCGGGCAGAGGATGCAACTTTTATCAACG	AGGTCTCTATAGATAAACCTTAGCGCTGGGA	4440
Db	4307	TAGCGGGCAGAGGATGCAACTTTTATCAACG	AGGTCTCTCTATAGATAAACCTTAGCGCTGGGA	4366
Qy	4441	TCATCTTTTGGACAACCTTTTCTGCCAAAATCT	TAGGTCGCAAAATCACATTCAATTCATACCAT	4500
Db	4367	TCATCTTTTGGACAACCTTTTCTGCCAAAATCT	TAGGTCGCAAAATCACATTCAATTCATACCAT	4426
Qy	4501	TATTGTACAACTTGAGCAAGTTGTCGATCAGCT	CCTCCATAATGGTCTCTGTAAACGGATG	4560

Db	4427	TATTGTACAACCTTGAGCAAGTTGTGCATCAGCTCCTCAAAATTGGTGCTCTGTGTAAACGGATG	4488
Qy	4561	ACTCAACTTGCACATTAACCTTGAAGCTCAGTCGATTCAGTGAACCTTGATCAGGTTGTGSCA	4620
Db	4487	ACTCAACTTGCACATTAACCTTGAAGCTCAGTCGATTCAGTGAACCTTGATCAGGTTGTGSCA	4546
Qy	4621	GCTGGTCAGCAGCATAGGGAACACGGCTTTTCTACCAAACTTCAAGGAATTTACAACCT	4680
Db	4547	GCTGGTCAGCAGCATAGGGAACACGGCTTTTCTACCAAACTTCAAGGAATTTACAACCT	4606
Qy	4681	CTGCAACAACCTTGCCTATGCAGGTAGCAAGGGAATGTCTACTTGAAGTCGGACAGTGAAG	4740
Db	4607	CTGCAACAACCTTGCCTATGCAGGTAGCAAGGGAATGTCTACTTGAAGTCGGACAGTGAAG	4666
Qy	4741	TGTAGTCTTGAGAAATCTGAAGCCGTATTTTATTTATCATAGTCAGTCAGTCATCAGGAGA	4800
Db	4667	TGTAGTCTTGAGAAATCTGAAGCCGTATTTTATTTATCATAGTCAGTCAGTCATCAGGAGA	4726
Qy	4801	TCCTCTACGCCGGACGATCGTGGCCGA-----	4828
Db	4727	TCCTCTACGCCGGACGATCGTGGCCGA-----	4786
Qy	4829	-----	4828
Db	4787	CTCGTGAAGAAGTGTGTCTCACTCATCACAGGCTGAATCGCCCATCATCCAGCCAGA	4846
Qy	4829	-----	4828
Db	4847	AAGTGAGGAGCCACGGTTGATGAGACTTTGTGTAGTGGACCACTTGGTGATTTTGA	4906
Qy	4829	-----	4828
Db	4907	ACTTTTGTCTTGCCACGGAACGGTCTGCGTGTGCGGAAGATCGTGATCTGATCCTTCA	4966
Qy	4829	-----	4828
Db	4967	ACTCAGCAAAAGTTCGATTTATTCAACAAAGCCCGTCCCGTCAAGTCAGCGTAATGCT	5026
Qy	4829	-----	4828
Db	5027	CTGCCAGTGTTCAAACCAATTACCAATTTCTGATTAGAAAAAATCATCGAGCATCAAAATG	5086
Qy	4829	-----	4828
Db	5087	AAACTGCAATTTATTCATATCAGGATTTATCAATACCATATTTTGTAAAAAGCCGTTTCTG	5146
Qy	4829	-----	4828
Db	5147	TAATGAGGAGAAAACTCACCGAGGCGTTCCATAGGATGGCAAGATCCTGGTATCGGTC	5206
Qy	4829	-----	4828
Db	5207	TGCGATTCGACTCGTCCAAATCAATACAACTTATTAATTTCCCTCGTCAAAAATAAG	5266
Qy	4829	-----	4828
Db	5267	GTTATCAAGTGAGAAATTCACATAGGTGACGACTGAATCCGGTGAGAAATGGCAAAAGCTT	5326
Qy	4829	-----	4828
Db	5327	ATGCAATTTCTTCAGACTTGTTCACAGGCCAGCCATACGGTCGTCATCAAAATCACT	5386
Qy	4829	-----	4828
Db	5387	CGCATCAACAAACCGTTATTCATTCGTGATTCGGCTGAGCGAGACGAATACGGATC	5446
Qy	4829	-----	4828
Db	5447	GCTGTTAAAGGACAATTACAAACAGGAATCGAATGCAACCGCGGACGAGACTGCCAG	5506
Qy	4829	-----	4828
Db	5507	CGCATCAACAAATTTTCACCTGAACTCAGGATATCTTCTAATACCTGGAATCGCTGTTT	5566

QY 4829 ----- 4828
Db 5567 CCGGGGATCGAGTGGTGAAGTAACCATGATCATCAGGAGTACGGATAAAATGCTTAT 5626
QY 4829 ----- 4828
Db 5627 GGTGGGAAGAGGAGTAATAATTCGGTCAGCGAGTTTAGTCTGACCATCTCATCTGTAAACATC 5686
QY 4829 ----- 4828
Db 5687 ATTGGCAACGCTACCTTTGGCCATGTTTCAGAAACAACTCTGCGCATCGGGCTTCCCAT 5746
QY 4829 ----- 4828
Db 5747 CAATCGATAGATTGTGCGACCTGATTGCCGACATTTATCGAGGCCATTTTATACCCATA 5806
QY 4829 ----- 4828
Db 5807 TAAATCAGCATCCATGTTGGAAATTTAATCGCGGCTCGAGCAAGACGTTTCCCGTTGAAT 5866
QY 4829 ----- 4828
Db 5867 ATGGCTATAACACCCCTTGTAATTAATGTAACGACAGTATTTATGTTTATGA 5926
QY 4829 ----- 4828
Db 5927 TGATATATTTTATCTTGTGCAATGTAATCATCAGAGATTTTGGACACAACTGGGCTTTC 5986
QY 4829 ----- 4879
Db 5987 CCCCCCCCCCTGAGTCCGATCACCGGCGCACAGTGCCTGCTGCGGCTATAT 6046
QY 4880 CGCGAGATCAACGATGGGGAAGATCGGGCTCGCCATCTTGGGCTCATAGAGCGCTGTTT 4939
Db 6047 CGCGAGATCAACGATGGGGAAGATCGGGCTCGCCATCTTGGGCTCATAGAGCGCTGTTT 6106
QY 4940 CGCGTGGGTATGTTGGCAGGCCCCGTCGGGGGAGTGTGGGGCCCATCTCTTGCA 4999
Db 6107 CGCGTGGGTATGTTGGCAGGCCCCGTCGGGGGAGTGTGGGGCCCATCTCTTGCA 6166
QY 5000 TGACCATTTCTTGGCGGGGGTCTCAACGGGCTCAACTACTAGTGGGCTGCTTCT 5059
Db 6167 TGACCATTTCTTGGCGGGGGTCTCAACGGGCTCAACTACTAGTGGGCTGCTTCT 6226
QY 5060 AATGACGAGTCCGATAGGAGAGCGTGCAGTATCTATGATTGGAAGTATGGGAATGT 5119
Db 6227 AATGACGAGTCCGATAGGAGAGCGTGCAGTATCTATGATTGGAAGTATGGGAATGT 6286
QY 5120 GATACCGCATTTCTCAGTGTCTGAGTCTCTATACAGATTAATGCCCACCTAAAGCAAC 5179
Db 6287 GATACCGCATTTCTCAGTGTCTGAGTCTCTATACAGATTAATGCCCACCTAAAGCAAC 6346
QY 5180 CGGAGGAGGATTTCTATGTAATTTCTCTGACTTTTGTGTCATCAGTACGAGTCTGAACTG 5239
Db 6347 CGGAGGAGGATTTCTATGTAATTTCTCTGACTTTTGTGTCATCAGTACGAGTCTGAACTG 6406
QY 5240 TGAGACTATCTCGGTTATGACGAGAAATGTCCTTTCTTGGAGACAGTAATAAGATGCC 5299
Db 6407 TGAGACTATCTCGGTTATGACGAGAAATGTCCTTTCTTGGAGACAGTAATAAGATGCC 6466
QY 5300 ACCAATAAGAAATCTTGTATCAGGAACAACTTCTTGTTCGAACTTTTTCGGTGCC 5359
Db 6467 ACCAATAAGAAATCTTGTATCAGGAACAACTTCTTGTTCGAACTTTTTCGGTGCC 6526
QY 5360 TTGAATATAAATGATAGTGAATGTTGGTAGGAATGGAGCGGCAAAATGCTTACC 5419
Db 6527 TTGAATATAAATGATAGTGAATGTTGGTAGGAATGGAGCGGCAAAATGCTTACC 6586
QY 5420 TTCTCGACCTTCAAGAGTATGAGGTTCTAGATCTGATGCAACTTTCAGTGACAAC 5479
Db 6587 TTCTCGACCTTCAAGAGTATGAGGTTCTAGATCTGATGCAACTTTCAGTGACAAC 6646

QY 5480 GTTGTATTTCTGTTCAACCATTCGAAATCAGAGAAATCAAAAGTTGTTGTCTACTATT 5539
Db 6647 GTTGTATTTCTGTTCAACCATTCGAAATCAGAGAAATCAAAAGTTGTTGTCTACTATT 6706
QY 5540 GATCCAAGCCAGTGGGCTTTGAAACTGACAATAGTGTGCTCGTGTGTTGAGTCTATCT 5599
Db 6707 GATCCAAGCCAGTGGGCTTTGAAACTGACAATAGTGTGCTCGTGTGTTGAGTCTATCT 6766
QY 5600 TGTATGATTAATCTAGTCTTTGATCTAAATTAATCTTGAAGCCAGGCGGATTAATACC 5659
Db 6767 TGTATGATTAATCTAGTCTTTGATCTAAATTAATCTTGAAGCCAGGCGGATTAATACC 6826
QY 5660 CAAATCTAAACTCTTTTAAACGTTAAAGGCAAGTATGCTGCTGTATTAACCC 5719
Db 6827 CAAATCTAAACTCTTTTAAACGTTAAAGGCAAGTATGCTGCTGTATTAACCC 6886
QY 5720 AAATCAGCTCTAGTCTGATCTCTCAATCACTTGAAGGCACTATCTGTTTGAAGAAAT 5779
Db 6887 AAATCAGCTCTAGTCTGATCTCTCAATCACTTGAAGGCACTATCTGTTTGAAGAAAT 6946
QY 5780 TTGCGGAGATCGGATATCGAGAAAGTACGCTGATTTTAAACGTTGAAATTTATCTCAA 5839
Db 6947 TTGCGGAGATCGGATATCGAGAAAGTACGCTGATTTTAAACGTTGAAATTTATCTCAA 7006
QY 5840 GATCTCTGCTCGCGGCTTTGCGTGAAGCGGATGCCGGAGCAGACAAAGCCGTCAGGGCGG 5899
Db 7007 GATCTCTGCTCGCGGCTTTGCGTGAAGCGGATGCCGGAGCAGACAAAGCCGTCAGGGCGG 7066
QY 5900 GAGACGCTCAAGCTTGTCTGTAAGCGGATGCCGGAGCAGACAAAGCCGTCAGGGCGG 5959
Db 7067 GAGACGCTCAAGCTTGTCTGTAAGCGGATGCCGGAGCAGACAAAGCCGTCAGGGCGG 7126
QY 5960 TCAGCGGCTTGGCGGGTGTGCGGGCGCAGCATGACCCAGTCACTAGTAGCATAGCGA 6019
Db 7127 TCAGCGGCTTGGCGGGTGTGCGGGCGCAGCATGACCCAGTCACTAGTAGCATAGCGA 7186
QY 6020 GTGTATATCTGCTTAACTATGCGGCATCAGAGCAGATTTGACTGAGAGTGCACCATATGC 6079
Db 7187 GTGTATATCTGCTTAACTATGCGGCATCAGAGCAGATTTGACTGAGAGTGCACCATATGC 7246
QY 6080 GGTGTAAATACCCACAGATGCTTAAGGAGAAATACCGCATCAGCGCTTTCGCTT 6139
Db 7247 GGTGTAAATACCCACAGATGCTTAAGGAGAAATACCGCATCAGCGCTTTCGCTT 7306
QY 6140 CCTGCTCACTGACTCGCTCGGCTCGGCTCGGCTGCGGCGAGCGGTATCAGCTCACT 6199
Db 7307 CCTGCTCACTGACTCGCTCGGCTCGGCTCGGCTGCGGCGAGCGGTATCAGCTCACT 7366
QY 6200 CAAAGCGGTAATACGTTATCCAGAAATCAGGGAATACCGAGAAAGACATGTGAG 6259
Db 7367 CAAAGCGGTAATACGTTATCCAGAAATCAGGGAATACCGAGAAAGACATGTGAG 7426
QY 6260 CAAAGCGGTAATACGTTATCCAGAAATCAGGGAATACCGAGAAAGACATGTGAG 6319
Db 7427 CAAAGCGGTAATACGTTATCCAGAAATCAGGGAATACCGAGAAAGACATGTGAG 7486
QY 6320 GGTCTCGGCGGCTTACCGGATACCGGATACCGGCTTTCGCTTTCGGAAGGTCGCG 6379
Db 7487 GGTCTCGGCGGCTTACCGGATACCGGATACCGGCTTTCGCTTTCGGAAGGTCGCG 7546
QY 6380 CGACGAGTATATAAGATACGAGCGTTCCTGCTGAGAGTCTCTGCTGCTCTCTG 6439
Db 7547 CGACGAGTATATAAGATACGAGCGTTCCTGCTGAGAGTCTCTGCTGCTCTCTG 7606
QY 6440 TTCCGAGCGGCTTACCGGATACCGGATACCGGCTTTCGCTTTCGGAAGGTCGCG 6499
Db 7607 TTCCGAGCGGCTTACCGGATACCGGATACCGGCTTTCGCTTTCGGAAGGTCGCG 7666
QY 6500 TTCTCTATGACTCAGCTGAGTATCTCAGTTCGCTGAGTTCGCTTTCGGAAGGTCGCG 6559
Db 7667 TTCTCTATGACTCAGCTGAGTATCTCAGTTCGCTGAGTTCGCTTTCGGAAGGTCGCG 7726
QY 6560 GCTGTGTGAGGAAACCCCGTTTACGCGGACCGCTTTCGCTTTCGGAAGGTCGCG 6619

Db 7727 GCTGTGTGTCAGGAACCCCGGTTGAGCCGCGCTGCGCTTATCCGGTAACATATCGTC 7786
Qy 6620 TTGAGTCCAAACCCGGTAAAGACACGACTTATCGCACTGGCAGCAGCCACTGGTAACAGGA 6679
Db 7787 TTGAGTCCAAACCCGGTAAAGACACGACTTATCGCACTGGCAGCAGCCACTGGTAACAGGA 7846
Qy 6680 TTGAGCAGCAGGATATAGCCGGTGTACAGAGTTCTTGAAGTGTGGGCTAACTACG 6739
Db 7847 TTGAGCAGCAGGATATAGCCGGTGTACAGAGTTCTTGAAGTGTGGGCTAACTACG 7906
Qy 6740 GCTACACTAGAGGACAGTATTTGATCTCGCTCTGCTCAAGCCAGTTACCTTCGGA 6799
Db 7907 GCTACACTAGAGGACAGTATTTGATCTCGCTCTGCTCAAGCCAGTTACCTTCGGA 7966
Qy 6800 AAAGAGTTGGTAGCTCTTGTATCCGGCAAAACCCAGCCGCTGGTAGCGGTGTTTTT 6859
Db 7967 AAAGAGTTGGTAGCTCTTGTATCCGGCAAAACCCAGCCGCTGGTAGCGGTGTTTTT 8026
Qy 6860 TTGCAAGCAGCAGATTAAGCGCAGAAAAGAGATCTCAAGAGATCCCTTGTATCTTT 6919
Db 8027 TTGCAAGCAGCAGATTAAGCGCAGAAAAGAGATCTCAAGAGATCCCTTGTATCTTT 8086
Qy 6920 CTACCGGCTCTGACGCTCAGTGGAAAGAACTCAAGTTAAGGGATTTTGGTCTAGAT 6979
Db 8087 CTACCGGCTCTGACGCTCAGTGGAAAGAACTCAAGTTAAGGGATTTTGGTCTAGAT 8146
Qy 6980 TATCAAAAGAGATCTTACCTAGATCTTTTAAATTAATAATGAAGTTTAAATCAATCT 7039
Db 8147 TATCAAAAGAGATCTTACCTAGATCTTTTAAATTAATAATGAAGTTTAAATCAATCT 8206
Qy 7040 AAGATATATAGTAACCTGGTCTGACAGTTACCAATGCTTAATCAGTCAGGACCTA 7099
Db 8207 AAGATATATAGTAACCTGGTCTGACAGTTACCAATGCTTAATCAGTCAGGACCTA 8266
Qy 7100 TCTCAGCAGTCTGTCTATTTGTTCCATCCATAGTTGCTGCTCCCGCTGGTAGATAA 7159
Db 8267 TCTCAGCAGTCTGTCTATTTGTTCCATCCATAGTTGCTGCTCCCGCTGGTAGATAA 8326
Qy 7160 CTACATAGCAGGAGGCTTACATCTGCGCCAGTCTGCAATGATATCCGCGAGCCAC 7219
Db 8327 CTACATAGCAGGAGGCTTACATCTGCGCCAGTCTGCAATGATATCCGCGAGCCAC 8386
Qy 7220 GCTCAGCGCTCCAGATTTATCAGCAATAAACAGCCAGCCGAGGCGCCGAGCGCAGAA 7279
Db 8387 GCTCAGCGCTCCAGATTTATCAGCAATAAACAGCCAGCCGAGGCGCCGAGCGCAGAA 8446
Qy 7280 GTGCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAGCTAGAG 7339
Db 8447 GTGCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAGCTAGAG 8506
Qy 7340 TAAGTAGTTCGCAAGTAAATAGTTTGGCAAGTTTGGCCATTTGCTGCGCATCGTGG 7399
Db 8507 TAAGTAGTTCGCAAGTAAATAGTTTGGCAAGTTTGGCCATTTGCTGCGCATCGTGG 8566
Qy 7400 TGTCAAGTCTCGTTCGTTGGTATGGCTTCAATCAGTCCGGTTCCTCAAGCAAGCGAG 7459
Db 8567 TGTCAAGTCTCGTTCGTTGGTATGGCTTCAATCAGTCCGGTTCCTCAAGCAAGCGAG 8626
Qy 7460 TTACATGATCCCGCTGTTGTCAAAAGCGGTTAGCTTCTCGTCTCGATCGTTG 7519
Db 8627 TTACATGATCCCGCTGTTGTCAAAAGCGGTTAGCTTCTCGTCTCGATCGTTG 8686
Qy 7520 TCAGAGTAGTGGCCGAGGTTTATCACTCATGTTATGGCAGCAGCTGCATAATCTC 7579
Db 8687 TCAGAGTAGTGGCCGAGGTTTATCACTCATGTTATGGCAGCAGCTGCATAATCTC 8746
Qy 7580 TTACTGTCTAGCATCCGTAAGATGCTTTTCTGTAGCTGGTGTACTCAACCAAGTCAT 7639
Db 8747 TTACTGTCTAGCATCCGTAAGATGCTTTTCTGTAGCTGGTGTACTCAACCAAGTCAT 8806
Qy 7640 TCTGAGATAGTGTATGCGGCAACCGAGTTGCTTGTGCGCGCTCAACACGGGATAATA 7699

Db 8807 TCTGAGATAGTGTATGCGGCGACCGAGTTGCTTCTGCCCGGCTCAACACGGGATAATA 8866
Qy 7700 CCGCGCCACATAGCAGAACTTTAAAGTCTCTCATATCGAAAGAGCTTTCTCGGGCGAA 7759
Db 8867 CCGCGCCACATAGCAGAACTTTAAAGTCTCTCATATCGAAAGAGCTTTCTCGGGCGAA 8926
Qy 7760 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCCGTGACCCCA 7819
Db 8927 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCCGTGACCCCA 8986
Qy 7820 ACTGATCTTTCAGCATCTTTTACTTTTCCACGAGGTTTCTGGGTGAGCAAAAACAGGAAGC 7879
Db 8987 ACTGATCTTTCAGCATCTTTTACTTTTCCACGAGGTTTCTGGGTGAGCAAAAACAGGAAGC 9046
Qy 7880 AAAATCCCGCAAAAAGGGAATAAGGGCGACACGGAATTTGGAATACTCATACTCTTCC 7939
Db 9047 AAAATCCCGCAAAAAGGGAATAAGGGCGACACGGAATTTGGAATACTCATACTCTTCC 9106
Qy 7940 TTTTTCATATTTATTCGAAGCAATTTATCAGGTTTATTTCTCATGACCGGATACATATTTG 7999
Db 9107 TTTTTCATATTTATTCGAAGCAATTTATCAGGTTTATTTCTCATGACCGGATACATATTTG 9166
Qy 8000 AATGTATTTAGAAAAATAAAATAGGGTTCGCGGCAATTTCCCGGAAAAGTCCAC 8059
Db 9167 AATGTATTTAGAAAAATAAAATAGGGTTCGCGGCAATTTCCCGGAAAAGTCCAC 9226
Qy 8060 CTGACGCTTAAGAAACCATTTATTCATGACATTAACCTTAAATAGCGGTATCACGA 8119
Db 9227 CTGACGCTTAAGAAACCATTTATTCATGACATTAACCTTAAATAGCGGTATCACGA 9286
Qy 8120 GGCCCTTTGCTCTCAAGAAATTAATTTCTCATGTTTACAGCTTATCATGATAAGCTGAC 8179
Db 9287 GGCCCTTTGCTCTCAAGAAATTAATTTCTCATGTTTACAGCTTATCATGATAAGCTGAC 9346
Qy 8180 TCATGTTGTTGTTGTAATAGACGAGATCCGGAACACTGGAATAAACAGTTATT 8239
Db 9347 TCATGTTGTTGTTGTAATAGACGAGATCCGGAACACTGGAATAAACAGTTATT 9406
Qy 8240 CG 8241
Db 9407 CG 9408

RESULT 5
AAF83654
ID AAF83654 standard; DNA; 9407 BP.
XX AC AAF83654;
XX DT 23-JUL-2001 (first entry)
XX DE pPIC9K-lunasin gene construct sequence.
XX KW Lunasin; cancer; soybean; variant; carcinogen; viral; oncogene; ds.
XX OS Glycine max.
XX FH Key Location/Qualifiers
XX FT 1222..1353
XX FT /*tag= a
XX FT /note= "lunasin DNA sequence"
XX FN WO200134808-A2.
XX PD 17-MAY-2001.
XX PF 12-NOV-2000; 2000WO-US031211.
XX PR 12-NOV-1999; 99US-0165334P.
XX PA (FILG-) FILGEN BIOSCIENCES INC.
XX PI Galvez AP;

XX WPI: 2001-343605/36.
 DR P-PSDB; AAB62613.
 XX
 PT Large scale production of lunasin, a cancer preventive peptide from
 PT soybean, by preparing lunasin gene constructs using pichia expression
 PT vectors, optimizing lunasin expression parameters and purifying lunasin.
 XX
 PS Disclosure; Page 65-67; 67pp; English.
 XX
 CC The invention provides a method of producing lunasin, a cancer preventive
 CC peptide from soybean, or its active variant by recombinant DNA technology
 CC in large quantities. The method comprises preparing lunasin gene
 CC constructs or its variants using protein expression vectors, optimizing
 CC lunasin expression parameters for large-scale production and performing a
 CC series of isolation and purification to obtain large quantities of
 CC biologically active recombinant lunasin peptide. The lunasin peptide
 CC exhibits inhibitory effects against malignant transformation of cells
 CC induced by chemical carcinogens and viral oncogenes. The present sequence
 CC represents the pPIC9K-lunasin gene construct sequence
 XX
 SQ Sequence 9407 BP; 2533 A; 2232 C; 2117 G; 2525 T; 0 U; 0 Other;

Query Match 81.0%; Score 6675.6; DB 4; Length 9407;
 Best Local Similarity 85.2%; Fred. No. 0;
 Matches 8082; Conservative 0; Mismatches 84; Indels 1316; Gaps 4;

1 AGATCTAATCAATCAAGAGCAAGAGTGAATGAATGAACCTTTTGGCCATCCGACATCCACAG 60
 1 AGATCTAATCAATCAAGAGCAAGAGTGAATGAATGAACCTTTTGGCCATCCGACATCCACAG 60
 61 GTCCATTCCTACATATAGTCCCAACGCAACAGAGGGGATACATAGCAGCAGCGT 120
 61 GTCCATTCCTACATATAGTCCCAACGCAACAGAGGGGATACATAGCAGCAGCGT 120
 121 TGCAACCGCAGACCTCCACTCTCTCTCTCTCAACACCCACTTTTGCCATCGAAAAACC 180
 121 TGCAACCGCAGACCTCCACTCTCTCTCTCTCAACACCCACTTTTGCCATCGAAAAACC 180
 181 AGCCGAGTTATGGGCTTGAATGAGAGTGGTCTCATTTCCAAATCTCTATAGGCTACTA 240
 181 AGCCGAGTTATGGGCTTGAATGAGAGTGGTCTCATTTCCAAATCTCTATAGGCTACTA 240
 241 ACACCATGACTTATAGGCTTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
 241 ACACCATGACTTATAGGCTTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
 301 TTTCCGAATGCAACAGCTCCGCAATATACACCGCAACATCACTCCAGATGAGGGCTTTCTG 360
 301 TTTCCGAATGCAACAGCTCCGCAATATACACCGCAACATCACTCCAGATGAGGGCTTTCTG 360
 361 AGTGTGGGGTCAATAGTTTATGTTTCCCAATAGGCGCCAAACTGACAGTTTAAACGCT 420
 361 AGTGTGGGGTCAATAGTTTATGTTTCCCAATAGGCGCCAAACTGACAGTTTAAACGCT 420
 421 GTCTTTGGAAACCTTAATATGCAAAAAAGCGTGTCTCATCCAGATGAACCTAAGTTTGGTTCG 480
 421 GTCTTTGGAAACCTTAATATGCAAAAAAGCGTGTCTCATCCAGATGAACCTAAGTTTGGTTCG 480
 481 TTGAATGCTAACGCCCAAGTGGTCAAAAGAAACTTCCAAAGTCCGCATACCGTTTGT 540
 481 TTGAATGCTAACGCCCAAGTGGTCAAAAGAAACTTCCAAAGTCCGCATACCGTTTGT 540
 541 CTGTGTTTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
 541 CTGTGTTTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
 601 CTCTATCGCTTGTGAACCCCGGTGCACTGTGCGGAAACGCAAAATGGGGAAACACCCGCT 660
 601 CTCTATCGCTTGTGAACCCCGGTGCACTGTGCGGAAACGCAAAATGGGGAAACACCCGCT 660
 661 TTTTGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720

661 TTTTGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 721 GCTGATAGCTTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTGAACAGCAAT 780
 721 GCTGATAGCTTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTGAACAGCAAT 780
 781 ATATATAACAGAGGAAGCTGCCCTGCTCTTAAACCTTTTATATCATCATTTATAGCTT 840
 781 ATATATAACAGAGGAAGCTGCCCTGCTCTTAAACCTTTTATATCATCATTTATAGCTT 840
 841 ACTTTTCAATATTCGAGCTGTTTCCAAATTTGAATGACAAAGCTTTTGAATTTAAACGCA 900
 841 ACTTTTCAATATTCGAGCTGTTTCCAAATTTGAATGACAAAGCTTTTGAATTTAAACGCA 900
 901 CAACTTGAGAGATCAAAAAACAATAATTTTGAAGGATCCAAACGATGAGATTTCTT 960
 901 CAACTTGAGAGATCAAAAAACAATAATTTTGAAGGATCCAAACGATGAGATTTCTT 960
 961 TCAATTTTACTGCGAGTTTATTCGAGCATCTCCCGCATTCCTCCAGTCAACACT 1020
 961 TCAATTTTACTGCGAGTTTATTCGAGCATCTCCCGCATTCCTCCAGTCAACACT 1020
 1021 ACAACAGAGATGAAACGGCACAATTCGGCTGAAGCTGTCTATCGGTTACTCAGATTTA 1080
 1021 ACAACAGAGATGAAACGGCACAATTCGGCTGAAGCTGTCTATCGGTTACTCAGATTTA 1080
 1081 GAAGGGGATTCGATGTTGCTGTTTGGCATTTTCCACAGCACAATAAAGGGTTATTG 1140
 1081 GAAGGGGATTCGATGTTGCTGTTTGGCATTTTCCACAGCACAATAAAGGGTTATTG 1140
 1141 TTTATAATATCTACTTATTCGAGCATTCGCTCTAAAGAGAGGGGTATCTCTCGAGAAA 1200
 1141 TTTATAATATCTACTTATTCGAGCATTCGCTCTAAAGAGAGGGGTATCTCTCGAGAAA 1200
 1201 AGAGAGGCTGAAGCCCGAGCCCGAGCCCGAGCCCGAGCCCGAGCCCGAGCCCGAGCC 1260
 1201 AGAGAGGCTGAAGCTTACGTAT--CCAAATGGCAGCACCAGCAAGATAGTGGCGCAAG 1260
 1261 GCGCGCGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
 1261 GCGCGCGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
 1258 GAGCTTCCAGGGGGTGAACCTCAGCCCTGCGAGAGACATCATCGAGAAATCCAGGC 1317
 1321 AGAGGTTGCTGTTTGTGATGACAGTGTCCGGGATTCGCGTGGTGTCCACCCCATGGCC 1380
 1318 GCGCGCGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1344
 1381 ATCGAGAACACTCAAGAGAGAAATGTCCTTCTAACTAGTGGGTAGAATTCCTTAGGG 1440
 1345 -----GACGACTAAGAAATTCCTTAGGG 1366
 1441 CGCGCGCGAAATTAATTCGCTTTAGACATGACTGTTTCCTCAGTTCAAGTTGGGCACTTACG 1500
 1367 CGCGCGCGAAATTAATTCGCTTTAGACATGACTGTTTCCTCAGTTCAAGTTGGGCACTTACG 1426
 1501 AGAAGACCGGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
 1427 AGAAGACCGGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1486
 1561 ATGACGGCTTCATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 1487 ATGACGGCTTCATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1546
 1621 GTCAATTTGTTTCTCTCGTACGAGCTTCTCTGATCAGCCTATCTCGAGCTGATGATA 1680
 1547 GTCAATTTGTTTCTCTCGTACGAGCTTCTCTGATCAGCCTATCTCGAGCTGATGATA 1606
 1681 TATCTTGTGTTGAGGGGTTTGGGAAATCATTTCCAGTTTGTGTTTCTTCTGTTATTTCCC 1740
 1607 TATCTTGTGTTGAGGGGTTTGGGAAATCATTTCCAGTTTGTGTTTCTTCTGTTATTTCCC 1666
 1741 ACTCTCTTTCAGAGTACAGAGATTAAGTGAAGCTTCTGTTTGTGCAAGCTTATCGATAA 1800
 1667 ACTCTCTTTCAGAGTACAGAGATTAAGTGAAGCTTCTGTTTGTGCAAGCTTATCGATAA 1726

3886 GCAACTTTTGTAGTTGCTCCAGATGTAGCACTTTTATACCAAAACCGGTGACGAGATT 3945
4021 GGTAGACTCCAGTTTGTGCTTATAGCTCCGGAATAGACTTTTGGACGAGTACACCA 4080
3946 GGTAGACTCCAGTTTGTGCTTATAGCTCCGGAATAGACTTTTGGACGAGTACACCA 4005
4081 GSCCAACGAGTAATAGAGAGTCCAGCAACAAAGTGTGAATAGACCATCGGGCGGT 4140
4006 GSCCAACGAGTAATAGAGAGTCCAGCAACAAAGTGTGAATAGACCATCGGGCGGT 4065
4141 CAGTAGTCAAGACGCCAACAAATTTTCACTGACAGGGAATTTTGTGACATCTTCAGAAA 4200
4066 CAGTAGTCAAGACGCCAACAAATTTTCACTGACAGGGAATTTTGTGACATCTTCAGAAA 4125
4201 GTTCGTATTTCAGTAGTCAATTCGCCGAGCATCAATTAATGGGATTATACCAAGCAACAG 4260
4126 GTTCGTATTTCAGTAGTCAATTCGCCGAGCATCAATTAATGGGATTATACCAAGCAACAG 4185
4261 TGGAGTCAATCTACCAACTTTTGGGCTCTCAGAAAAGCATTAACAGTCTTACTACGC 4320
4186 TGGAGTCAATCTACCAACTTTTGGGCTCTCAGAAAAGCATTAACAGTCTTACTACGC 4245
4321 CATTAGTGAATCTTTCAATCGCCAGTGGAGAGAAAAGGCACAGCGATCTAGCAT 4380
4246 CATTAGTGAATCTTTCAATCGCCAGTGGAGAGAAAAGGCACAGCGATCTAGCAT 4305
4381 TAGGGGCAAGGATGCACTTTATCAACAGGCTCCTATAGATTAACCTAGCGCTGGGA 4440
4306 TAGGGGCAAGGATGCACTTTATCAACAGGCTCCTATAGATTAACCTAGCGCTGGGA 4365
4441 TCATCTTTGGCAACTCTTTCTGCAAACTTAGGTCCAAATCACTTCAATGATACCAT 4500
4366 TCATCTTTGGCAACTCTTTCTGCAAACTTAGGTCCAAATCACTTCAATGATACCAT 4425
4501 TATTGTACAACTTGAGCAAGTTGTGATCAGCTCTCAAAATGGTCTCTGTAAACGATG 4560
4426 TATTGTACAACTTGAGCAAGTTGTGATCAGCTCTCAAAATGGTCTCTGTAAACGATG 4485
4561 ACTCACTTTGCATTAATCTTCAAGTCACTGATGAGTGAATGATCACTGATGTCGA 4620
4486 ACTCACTTTGCATTAATCTTCAAGTCACTGATGAGTGAATGATCACTGATGTCGA 4545
4621 GCTGCTCAGCAGATAGGGAACACGGCTTTTCTACCAAACTCAAGGAATTTACAACT 4680
4546 GCTGCTCAGCAGATAGGGAACACGGCTTTTCTACCAAACTCAAGGAATTTACAACT 4605
4681 CTGCAACACTTGGGTATGAGTAGCAAGGGAATGTATCACTTGAAGTCGACAGTGA 4740
4606 CTGCAACACTTGGGTATGAGTAGCAAGGGAATGTATCACTTGAAGTCGACAGTGA 4665
4741 TGTAGCTTTGAGAAATCTTGAAGCGGTATTTTATATCACTGAGTCACTCAGGAGA 4800
4666 TGTAGCTTTGAGAAATCTTGAAGCGGTATTTTATATCACTGAGTCACTCAGGAGA 4725
4801 TCTCTACCGCGGACGCATCTGTGGCGA----- 4828
4726 TCTCTACCGCGGACGCATCTGTGGCGA----- 4785
4829 ----- 4828
4786 CTGCTGAAGAGTGTGTGCTGACTCATACCGGCTGATTCGCCCATCATCCAGCCAGA 4845
4829 ----- 4828
4846 AAGTAGGAGGACCGGTTGATGAGAGCTTTGTTGTAGGTGGACGAGTTGATTTTGA 4905
4829 ----- 4828
4906 ACTTTTGTTCGCCAGGAACGGTCTGCGTTGTTCGGGAAGATGCGTGATCTGATCCCTCA 4965
4829 ----- 4828

4966 ACTCAGCAAAAGTTTCGATTTTATCAACAAAGCCGCGTCCCGTCAAGTCAGCGTAATGCT 5025
4829 ----- 4828
5026 CTGCCAGTGTACACCAATTAACCAATTTCTGATTAGAAAACTCATCGAGCATCAAAATG 5085
4829 ----- 4828
5086 AAACCTGCAATTTTATCATATCAGGATTATCAATACCATATTTTGA AAAAGCGGTTTCG 5145
4829 ----- 4828
5146 TAATGAAGAGGAAAACTCACCAGGAGTTCATAGGATGGCAAGATCTCGTATCGGTC 5205
4829 ----- 4828
5206 TGGATTCCGACTCGTCCAAACATCAATCAACCTATTAATTTCCCTCGTCAAAAATAG 5265
4829 ----- 4828
5266 GTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAAATGGCAAAAGCTT 5325
4829 ----- 4828
5326 ATGCATTTCTTTCCAGACTTTGTTCAACAGCCAGCCATTAACGCTCGTCAATAATCACT 5385
4829 ----- 4828
5386 CGCATCAACCAACCGTTTATTCATTCGTGATTGGCGCTGAGCGAGACGAAATACGGATC 5445
4829 ----- 4828
5446 GCTGTTAAAAAGGACAATTACAAACAGGAATCGAATGCAACCGCGCAGGAAACACTGCCAG 5505
4829 ----- 4828
5506 CGCATCAACAAATTTTCACTGATCAGGATATTTCTTAATACCTGGAATGCTGTTT 5565
4829 ----- 4828
5566 CCCGGGATCGCAGTGGTGAATAACCATCATCAGGATACGGAATAAAATGCTTGAT 5625
4829 ----- 4828
5626 GGTCCGAAGAGGATAAATTCGCTCAGCCAGTTTGTAGTCTGACCATCTCATCTGTAACATC 5685
4829 ----- 4828
5686 ATTGGCAACGCTACTTTTGCATGTTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCAT 5745
4829 ----- 4828
5746 CAATCGATAGATTGCGACACTGATTTCGCCGACATTAATCGCGACCCATTATACCCATA 5805
4829 ----- 4828
5806 TAAATCAGCATCCATGTTTGAATTTAATCGCGCTCTGAGCAAGAGGTTTCCCGTTGAAT 5865
4829 ----- 4828
5866 ATGGCTCATAAACCCCTTGTATTACTGTTTATGTAAAGCAGACAGTTTATTGTTTATGA 5925
4829 ----- 4828
5926 TGATATATTTTATCTTGTGCAATGTAAATCAGAGATTTTGAACACAACTGGCTTTC 5985
4829 ----- CCTCAGGTCCGCATCACCAGCGCCACAGGTGCGGTTGCTGGCGCTTATAT 4879
5986 CCCCCCCCCCTCGAGTCCGCATCACCAGCGCCACAGGTGCGGTTGCTGGCGCTTATAT 6045
4880 CCGCGACATCAGGATGGGAGATCGGCTCGCACTTCGGCTCATGAGCGCTTGT 4939
6046 CCGCGACATCAGGATGGGAGATCGGCTCGCACTTCGGCTCATGAGCGCTTGT 6105

Db 8266 TCTCAGCATCTGTCTATTTTCGTTTATCATCATAGTTGCTGACTCCCGCTCGTGTGATATAA 8325
Qy 7160 CTACGATACGGAGGCTTACCATCTGGCCCCAGTGTCTGCAATGATACCGGAGACCCAC 7219
Db 8326 CTACGATACGGAGGCTTACCATCTGGCCCCAGTGTCTGCAATGATACCGGAGACCCAC 8385
Qy 7220 GCTCACCCTCCAGATTTATCAGCAATAAACCAGCCGCGGAGGCGCGGAGGCGAGAA 7279
Db 8386 GCTCACCCTCCAGATTTATCAGCAATAAACCAGCCGCGGAGGCGGAGGCGAGAA 8445
Qy 7280 GTGGTCTCTGCAACTTTATCCGCTCCATCCAGTCTTATTAATTTGTTCCGCGGAAGCTAGAG 7339
Db 8446 GTGGTCTCTGCAACTTTATCCGCTCCATCCAGTCTTATTAATTTGTTCCGCGGAAGCTAGAG 8505
Qy 7340 TAAGTAGTTCGCGAGTTAATAGTTTTCGCAACAGTTGTTGCCATTTGCTGCAGGCATCTGG 7399
Db 8506 TAAGTAGTTCGCGAGTTAATAGTTTTCGCAACAGTTGTTGCCATTTGCTGCAGGCATCTGG 8565
Qy 7400 TGTCAAGTTCGCTGCTGTTGTTGATGCTTCAATTCAGCTCCGCTTCCCAACGATCAAGGCGAG 7459
Db 8566 TGTCAAGTTCGCTGCTGTTGTTGATGCTTCAATTCAGCTCCGCTTCCCAACGATCAAGGCGAG 8625
Qy 7460 TTACATGATCCCCCATGTTTGTCAAAAAGAGGTTAGTCTCTTCGCTCCGATCGTTG 7519
Db 8626 TTACATGATCCCCCATGTTTGTCAAAAAGAGGTTAGTCTCTTCGCTCCGATCGTTG 8685
Qy 7520 TCAGAAGTAAGTTGCGCGAGTGTATCATCATGTTTATGCGAGCACTGCATAATCTC 7579
Db 8686 TCAGAAGTAAGTTGCGCGAGTGTATCATCATGTTTATGCGAGCACTGCATAATCTC 8745
Qy 7580 TTACTGTATGCGATCCGTAAGATCTTTTCTGTGACTGTGTGACTTCAACCAAGTCAAT 7639
Db 8746 TTACTGTATGCGATCCGTAAGATCTTTTCTGTGACTGTGTGACTTCAACCAAGTCAAT 8805
Qy 7640 TCTGAGATAGTGTATGCGCGAGTGTCTTCTGCGCGGCTCAACCGGATATAA 7699
Db 8806 TCTGAGATAGTGTATGCGCGAGTGTCTTCTGCGCGGCTCAACCGGATATAA 8865
Qy 7700 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCTGGAAGAACTTCTTCGCGGCGAA 7759
Db 8866 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCTGGAAGAACTTCTTCGCGGCGAA 8925
Qy 7760 AACTCTCAGGATCTTACCGCTGTGAGATCCAGTTCAGTGTACCCACTCGTGACCCA 7819
Db 8926 AACTCTCAGGATCTTACCGCTGTGAGATCCAGTTCAGTGTACCCACTCGTGACCCA 8985
Qy 7820 ACTGATCTTCAGCATCTTTTACTTTTCAACGAGTCTTCTGCGTGAAGAAACAGGAAGGC 7879
Db 8986 ACTGATCTTCAGCATCTTTTACTTTTCAACGAGTCTTCTGCGTGAAGAAACAGGAAGGC 9045
Qy 7880 AAATGCCGCAAAAAGGGAATAAGGCGACACGGAATGTTGAATCTCATACTCTTCC 7939
Db 9046 AAATGCCGCAAAAAGGGAATAAGGCGACACGGAATGTTGAATCTCATACTCTTCC 9105
Qy 7940 TTTTTCATATTTTGAAGCATTTTATCAGGTTTATGTTCTCATGAGCGGATACATATTG 7999
Db 9106 TTTTTCATATTTTGAAGCATTTTATCAGGTTTATGTTCTCATGAGCGGATACATATTG 9165
Qy 8000 AATGTATTTAGAAAATAAACAATAGGGTTTCGCGCAATTTCCCGCAAAAGTGCCAC 8059
Db 9166 AATGTATTTAGAAAATAAACAATAGGGTTTCGCGCAATTTCCCGCAAAAGTGCCAC 9225
Qy 8060 CTGAGCTCTAAGAACCATTTATCATCATTAACCTATAAATAAGCGGTATCACGA 8119
Db 9226 CTGAGCTCTAAGAACCATTTATCATCATTAACCTATAAATAAGCGGTATCACGA 9285
Qy 8120 GGCCCTTTTCTGTTTCAAGAAATTAATCTCATGTTTGAAGCTTATCATGATTAAGCTGAC 8179
Db 9286 GGCCCTTTTCTGTTTCAAGAAATTAATCTCATGTTTGAAGCTTATCATGATTAAGCTGAC 9345
Qy 8180 TCATGTTGGTATTTCTGAATAGCGGATCGGGAACACTGAAATAACAGTTATATT 8239

Db 9346 TCAATGTTGTTATGTTGAAATAGCGAGATCGGGAACACTGAAATAACAGTTATATT 9405
Qy 8240 CG 8241
Db 9406 CG 9407
RESULT 6
AAI77818
ID AAT77818 standard; DNA; 9191 BP.
XX AAT77818;
AC AC
XX 25-MAR-2003 (revised)
DT 03-MAR-1998 (first entry)
XX
XX Neisseria meningitidis class 3 mature porin gene in pnv318.
KW Porin protein; porB gene; outer membrane protein; MB3; vaccine;
KW meningitis; Pichia pastoris; pnv318; vector; pHIL-S1; protein secretion;
ds.
XX Neisseria meningitidis serogroup B; strain 8765 (B:15:P1.
OS 3).
OS Synthetic.
OS Chimeric.
XX
PH Key Location/Qualifiers
FT CDS 1011..1949
FT /tag= a
FT /product= "mature porin"
FT mutation 1014..1016
FT /tag= b
FT /note= "GTT to GTC alteration of codon usage"
FT mutation 1017..1019
FT /tag= c
FT /note= "ACC to ACT alteration of codon usage"
FT mutation 1020..1022
FT /tag= d
FT /note= "CTG to TTG alteration of codon usage"
FT mutation 1026..1028
FT /tag= e
FT /note= "GGC to GGT alteration of codon usage"
FT mutation 1029..1031
FT /tag= f
FT /note= "ACC to ACT alteration of codon usage"
FT mutation 1032..1034
FT /tag= g
FT /note= "ATC to ATT alteration of codon usage"
FT mutation 1035..1037
FT /tag= h
FT /note= "AAA to AAG alteration of codon usage"
FT mutation 1038..1040
FT /tag= i
FT /note= "GCC to GCT alteration of codon usage"
FT mutation 1041..1043
FT /tag= j
FT /note= "GGC to GGT alteration of codon usage"
FT mutation 1044..1046
FT /tag= k
FT /note= "GTA to GTT alteration of codon usage"
FT mutation 1047..1049
FT /tag= l
FT /note= "GAA to GAG alteration of codon usage"
XX
PN WO9728273-A1.
XX
PD 07-AUG-1997.
XX
PP 31-JAN-1997; 97MO-US001687.
XX
PR 01-FEB-1996; 96US-0010972P.
PR 13-JUN-1996; 96US-0020440P.


```
XX PA (NAVA-) NORTH AMERICAN VACCINE INC.
XX PI Tai JY, Donets M, Wang M, Liang S, Polvinobodnar M, Minetti CASA;
XX PI Michon F;
XX DR WPI; 1997-402625/37.
XX DR P-PSDB; AA21744.
XX PT Production of meningococcal group B porin proteins in yeast - by
XX PT extraction of a porin protein, or porin protein fused to a yeast
XX PT secretion signal peptide.
XX PS Example 9; Fig 17A-C; 138pp; English.
XX CC This DNA sequence, designated clone pnv318, comprises a modified coding
XX CC sequence for the mature class 3 porin (MB3) protein (see AA21744) of
XX CC group B Neisseria meningitidis inserted into the XhoI-BamHI sites of
XX CC expression vector pMIL-SiD2, i.e. directly downstream from the PHO1
XX CC leader peptide, in frame with the secretion signal open reading frame for
XX CC secretion of expressed protein. The 5' end of the MB3 gene is optimised
XX CC for yeast codon usage (see also AA76738) to improve translation
XX CC efficiency in yeast host cells. MB3 was expressed in Pichia pastoris
XX CC transformants at a level of 300-400 mg/l. Methods are claimed for the
XX CC high level expression and purification of meningococcal group B porin
XX CC proteins from yeast cells. The porin proteins are used to provide an
XX CC immune response in mammals to N. meningitidis (claimed). (Updated on 25-
XX CC MAR-2003 to correct PI field.)
XX SQ Sequence 9191 BP; 2406 A; 2221 C; 2087 G; 2477 T; 0 U; 0 Other;

Query Match 79.9%; Score 6582; DB 2; Length 9191;
Best Local Similarity 85.9%; Pred. No. 0;
Matches 7911; Conservative 0; Mismatches 315; Indels 980; Gaps 16;

QY 1 AGATCTAATCAATCAAGAGCAAGAGGTTGAATGAACCTTTTGGCATCCGACATCCACAG 60
DB 1 AGATCTAATCAATCAAGAGCAAGAGGTTGAATGAACCTTTTGGCATCCGACATCCACAG 60

QY 61 GTCCATTCTCACATAAGTGCCAAACGCAAGAGGGGATACACTAGCAGCAGACCGT 120
DB 61 GTCCATTCTCACATAAGTGCCAAACGCAAGAGGGGATACACTAGCAGCAGACCGT 120

QY 121 TGCACACGAGACCTCCACTCTCTCTCTCACACCCACTTTTGGCATCCGAAACACC 180
DB 121 TGCACACGAGACCTCCACTCTCTCTCTCACACCCACTTTTGGCATCCGAAACACC 180

QY 181 AGCCAGATTATTGGGCTTGATTGGAGCTCGCTCAATTCCTCTTATTAGGCTACTA 240
DB 181 AGCCAGATTATTGGGCTTGATTGGAGCTCGCTCAATTCCTCTTATTAGGCTACTA 240

QY 241 ACACCATGACTTTATTAGGCTTGCTATCTGCCCCCTGGGAGGTTCAATGTTGTTTA 300
DB 241 ACACCATGACTTTATTAGGCTTGCTATCTGCCCCCTGGGAGG-TCAATGTTGTTTA 299

QY 301 TTTCCGAATGCAACAACTCCGATTAACCCGAAACATCACTCCAGATGAGGGCTTCTG 360
DB 301 TTTCCGAATGCAACAACTCCGATTAACCCGAAACATCACTCCAGATGAGGGCTTCTG 359

QY 361 AGTGTGGGTCAAATAGTTTCAATGTTCCCAATGGCCCAAAATCGACAGTTTAAACGCT 420
DB 361 AGTGTGGGTCAAATAGTTTCAATGTTCCCAATGGCCCAAAATCGACAGTTTAAACGCT 418

QY 421 GTCTTGGAACTTAATATGACAAAGCGTGATCTCATCAAGATGAATCAAGTTGGTTTCG 480
DB 419 GTCTTGGAACTTAATATGACAAAGCGTGATCTCATCAAGATGAATCAAGTTGGTTTCG 478

QY 481 TTGAATGCTAACGGCCAGTTGGTCAAAAAGAAATCTTCAAAGTCGCCATACCGTTTGT 540
DB 479 TTGAATGCTAACGGCCAGTTGGTCAAAAAGAAATCTTCAAAGTCGCCATACCGTTTGT 538

QY 541 CTTGTTTGTATGATGACGATGCTCAAAATATATCTCATTAATGCTTAGCCGAGTCT 600
DB 541 CTTGTTTGTATGATGACGATGCTCAAAATATATCTCATTAATGCTTAGCCGAGTCT 600

539 CTTGTTTGTATGATGACGATGCTCAAAATATATCTCATTAATGCTTAGCCGAGTCT 598
601 CTTGTTTGTATGATGACGATGCTCAAAATATATCTCATTAATGCTTAGCCGAGTCT 659
599 CTTGTTTGTATGATGACGATGCTCAAAATATATCTCATTAATGCTTAGCCGAGTCT 658
660 TTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 718
659 TTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 718
719 CTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 777
719 CTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 778
778 AATATATAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 837
779 AATATATAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 838
838 CTTTACCTTTCATTAATGCGAGTGGTCCCAATTCGACAGCTTTTGAATTTAAACGCTTTTAA 897
839 CTTTACCTTTCATTAATGCGAGTGGTCCCAATTCGACAGCTTTTGAATTTAAACGCTTTTAA 898
898 CGACAACTTCGAGAGAGATCAAAAACCAATTAATTCGAA-----GGATCCAAACGAT 950
899 CGACAACTTCGAGAGAGATCAAAAACCAATTAATTCGAAACGATGTTCTCTCCAATTTT 958
951 GAGATTTCTTCAATTTTACTGCAATTTTATTCGAGCATCTTCGCAATTAGCTGCTCC 1010
959 GTCTTGGAAATTAATTTAGCTTTGGCTACTTTTGAATCTGTCTCTCGAGACGCTC 1018
1011 AGTCAACACTACAAACAGAGAGAGAG-----1033
1019 TTTGTTAGGCTACTTAATTAAGCTGGTGTGAGACTTCCGCTCTGTATTTTACCAGAACGG 1078
1034 --AAACGGCACAATTTCCGCTGAACTGTCTGTTACTGTTACTGAGATTTAGAGGGGATTT 1091
1079 CCAAGTTACTGAAAGTTTACAACTACCGCTACCGCTCGTTGATTTGGTTTGAATAATCGGCTT 1138
1092 CGATG-----TTGCTGTTTGGCTTTTCCAAACAGACACAAATAACG 1132
1139 CAAAGGCCAAGAGAGACCTGGTAAACGCTGAAAGCCATTTGGCAGGTGAGCAAAAAGC 1198
1133 GGTATTATTGTTATA-----AATACTACTATTGCGCAGATTTG 1168
1199 ATCTATCCGCGTACTGACTCCGGTTGGGCAACCGCAATCTTCTATCGGCTTTGAAGG 1258
1169 CTGCTAAAGAGAGGGGTATCTCTCGAAGAAAGAGAG-----1206
1259 CGGCTTCGGTAAATTCGCGTCTGCTGCTTTGAAACAGCGCTCTGAAAGACACCGGCGACAT 1318
1207 -----GCTGAAGCCGAGG 1220
1319 CAATCTTGGGATAGCAAAAGCGACTATTGTTGGTGTAAACAAATTTGCCGAAACCGAGGC 1378
1221 CCAGGCCAGGCCAGGCCAGGAGAAACATGATCATGTCGCCCCCGGGAGAGGATATAA 1280
1379 ACGCTCATTTCCGCTACGCTACGATTTCTCCGAAATTTGCGGCTTCAGCGGCGAGCTACA 1438
1281 TTGTGGCTTCCCGGCTGTCACCGCCAGGAGTGCACGGAG-----1320
1439 ATACGCGCTTAAACGACAAATGACGAGAGACATAACAGCGAATCTTACACCGCGGTTCAA 1498
1321 -----AGAGGTTGCTGTTTGTGACAGTGTCCG-----1349
1499 CTACAAAACCGTGGCTTCTTCTGTCGAAATATGCGGCTGCTATAAAAAGACATCATCAAGT 1558
1350 -----1349
1559 GCAAGAGGGCTTGAAATATTGAGAAATACAGATTTCCGTTTGGTTCAGCGGTTACGACAA 1618
1350 -----1349
1619 TGATGCCCTGTACGCTTCCGTACGAGCAACAGAGCGGAACTGACTGATGCTTC 1678
```

QY 1350 ----- 1349
Db 1679 CAATTGCGACAACTCTCAAAACCGAAGTTGCGCTACCTTGGCATACCGCTTCGGCAACGT 1738
QY 1350 ----- 1349
Db 1739 AACGCCCGAGTTTCTTAAGCCCAACGCTTCAAGGTTTGGTTGATGATGACAGACATAGG 1798
QY 1350 -----GGGATTCGCCGT 1360
Db 1799 CAACGAATACGACCAAGTGTGCTGCGTGGGAATACGACTTCTCCAAACGCACTTCTGC 1858
QY 1361 GGTGCTTCCACCCCATGCGCATACGAGAACACTCAAGAAAGAAATGTCCTTCTAATAG 1420
Db 1859 CTTGGTTTCTGCCGTTGCTTGCAGAAAGGCAAGGCGAATAAATAATTCGTAGCGACTGC 1918
QY 1421 TGGCGTAGAATTCCTAGGGCGCGCGAATTAATTCGCTTAGACATGACTGTCTCTCA 1480
Db 1919 CGCGGTGTGCTGCGCCCAAAATCTAATCTGGATCCTTAGACATGACTGTCTCTCA 1978
QY 1481 GTTCAAGTTGGGCACTTACGAGAAAGCCGCTTTCCTAGATCTAATCAAGAGAGATGTC 1540
Db 1979 GTTCAAGTTGGGCA-TTACGAGAAAGCCGCTTTCCTAGATCTAATCAAGAGAGATGTC 2037
QY 1541 GAATGCCATTTGCGTGGAGATGAGGCTTCATTTTGTATCTTTTGTATCTTTTGTAACTTA 1600
Db 2038 GAATGCCATTTGCGTGGAGATGAGGCTTCATTTTGTATCTTTTGTATCTTTTGTAACTTA 2097
QY 1601 TATAGTATAGGATTTTTTTTCTCAATTTTGTCTTCTGACGAGCTTCTCTGATCAG 1660
Db 2098 TATAGTATAGGATTTTTTTTGTCAATTTTGTCTTCTGACGAGCTTCTCTGATCAG 2157
QY 1661 CCTATCTCGAGCTGATGAATATCTTGTGTAGGGTTTGGGAAATCAATTCAGATTGA 1720
Db 2158 CCTATCTCGAGCTGATGAATATCTTGTGTAGGGTTTGGGAAATCAATTCAGATTGA 2217
QY 1721 TGTCTTCTGCTGATTTCCCACTCTCTTACAGATACAGAAAGATTAAGTACAGAGTTTCTG 1780
Db 2218 TGTCTTCTGCTGATTTCCCACTCTCTTACAGATACAGAAAGATTAAGTACAGAGTTTCTG 2277
QY 1781 TTGTGCAAGCTTATCGATTAAGCTTAAATGCGGTAGTTTATCAAGTTAAATTTGCTAACGC 1840
Db 2278 TTGTGCAAGCTTATCGATTAAGCTTAAATGCGGTAGTTTATCAAGTTAAATTTGCTAACGC 2337
QY 1841 AGTCAGGACCGTGTATGAATCTAACAATGCGTCTATCGTCTATCGTCTCGGACCGTCAAC 1900
Db 2338 AGTCAGGACCGTGTATGAATCTAACAATGCGTCTATCGTCTATCGTCTCGGACCGTCAAC 2397
QY 1901 CTGGATGCTGTAGGCATAGGCTTGGTTATGCGGTACTGCGGCGCTCTTTCGGGATATC 1960
Db 2398 CTGGATGCTGTAGGCATAGGCTTGGTTATGCGGTACTGCGGCGCTCTTTCGGGATATC 2457
QY 1961 GTCCATTCGACAGCATCGGACGATCACTATGCGGTGCTGTAGCGCTATATGCTGTGATG 2020
Db 2458 GTCCATTCGACAGCATCGGACGATCACTATGCGGTGCTGTAGCGCTATATGCTGTGATG 2517
QY 2021 CAATTTCTATGCGCACCGCTTCTCGGAGCACTGTCGACCGCTTTGCGCGCGCCCAAGTC 2080
Db 2518 CAATTTCTATGCGCACCGCTTCTCGGAGCACTGTCGACCGCTTTGCGCGCGCCCAAGTC 2577
QY 2081 CTGCTCGCTTCTGCTACTTGGAGCACTATCGACTACGGGATCATGGGACCAACACCGCTC 2140
Db 2578 CTGCTCGCTTCTGCTACTTGGAGCACTATCGACTACGGGATCATGGGACCAACACCGCTC 2637
QY 2141 CTGTGGATCTATCGAATCTAATGTAGTTAAATCTCTAATTAATTAATTAATTAATTAATTA 2200
Db 2638 CTGTGGATCTATCGAATCTAATGTAGTTAAATCTCTAATTAATTAATTAATTAATTAATTA 2697
QY 2201 TTTCTCCATAGAACCTTAAAGCATTTGCGGTGAGCATCTAGACCTTCAACAGCAGCCAG 2260
Db 2698 TTTCTCCATAGAACCTTAAAGCATTTGCGGTGAGCATCTAGACCTTCAACAGCAGCCAG 2757

QY 2261 ATCCATCACTGCTTGGCCCAATATGTTTCAGTCCCTCAGAGTTACGCTTGTGAGGTGAT 2320
Db 2758 ATCCATCACTGCTTGGCCCAATATGTTTCAGTCCCTCAGAGTTACGCTTGTGAGGTGAT 2817
QY 2321 GAACTTCTGGAAGGTTGAGTGAATTAACCTCCGCTGTATTGACGGCATATCGTACGTTGG 2380
Db 2818 GAACTTCTGGAAGGTTGAGTGAATTAACCTCCGCTGTATTGACGGCATATCGTACGTTGG 2877
QY 2381 CAAAGTGTGTTGTTACCGGAGGATATCTCCAACTCTCTCGAGAGTAGGACCAAC 2440
Db 2878 CAAAGTGTGTTGTTACCGGAGGATATCTCCAACTCTCTCGAGAGTAGGACCAAC 2937
QY 2441 AAACACAGATCCAGGCTGTGTGATCAATCAATCAATCAATCAATCAATCAATCAATCA 2500
Db 2938 AAACACAGATCCAGGCTGTGTGATCAATCAATCAATCAATCAATCAATCAATCAATCA 2997
QY 2501 GATCAAGTGTTCAGGAGGCTACTGATTGGAATTTCCAAAGCCTGCTCGTAGGTTGCAAC 2560
Db 2998 GATCAAGTGTTCAGGAGGCTACTGATTGGAATTTCCAAAGCCTGCTCGTAGGTTGCAAC 3057
QY 2561 CGATAGGCTGTAGAGTGTGCAATCACTTGGTACAAATTTCAACCTTGGCAACTGCAC 2620
Db 3058 CGATAGGCTGTAGAGTGTGCAATCACTTGGTACAAATTTCAACCTTGGCAACTGCAC 3117
QY 2621 AGCTTGGTTGGAACAGCATCTTCAATTTCTGGCAAGCTCCTTGTCTGTCTATCGACAGC 2680
Db 3118 AGCTTGGTTGGAACAGCATCTTCAATTTCTGGCAAGCTCCTTGTCTGTCTATCGACAGC 3177
QY 2681 GAAAGATACCTTGGGATCAATACCAATGTTGAGTGTGAGACAGAGGCTCTGAGGCAAC 2740
Db 3178 GAAAGATACCTTGGGATCAATACCAATGTTGAGTGTGAGACAGAGGCTCTGAGGCAAC 3237
QY 2741 GAAATCTGGATCAGGCTATTTTATCAGCAATTAATAGAACTTTCAGAAAGCCCGCAGGAGCAT 2800
Db 3238 GAAATCTGGATCAGGCTATTTTATCAGCAATTAATAGAACTTTCAGAAAGCCCGCAGGAGCAT 3297
QY 2801 GTCAATATACACAGGCTGTATGTTGATTTTGAACCAATCATCTTGGAGCAGTAAACGAA 2860
Db 3298 GTCAATATACACAGGCTGTATGTTGATTTTGAACCAATCATCTTGGAGCAGTAAACGAA 3357
QY 2861 CTGCTTCTCGAACCAATATTTTGTACACTTAGGAACAGTTTCTGTTCCGTAAAGCAT 2920
Db 3358 CTGCTTCTCGAACCAATATTTTGTACACTTAGGAACAGTTTCTGTTCCGTAAAGCAT 3417
QY 2921 AGCAGCTACTGCTCGGCGCTCTCTGTAGCAAGATACACTTAGCACCAACTTGTGGGC 2980
Db 3418 AGCAGCTACTGCTCGGCGCTCTCTGTAGCAAGATACACTTAGCACCAACTTGTGGGC 3477
QY 2981 AACGTAGATGACTTCTGGGTAAAGGTACCACTCTTCTAGGTGAGATGCAAAACAAAT 3040
Db 3478 AACGTAGATGACTTCTGGGTAAAGGTACCACTCTTCTAGGTGAGATGCAAAACAAAT 3537
QY 3041 TTCTTTTCAACACAGCAAACTTTTGGCAGGAACCCAGCATCAGGGAAGTGGAGGAGCAAT 3100
Db 3538 TTCTTTTCAACACAGCAAACTTTTGGCAGGAACCCAGCATCAGGGAAGTGGAGGAGCAAT 3597
QY 3101 TCGGTTTCCACAGGAATATAGAGCCCACTTCTCAATAGGTCTTGCAGAAACAGGAGCA 3160
Db 3598 TCGGTTTCCACAGGAATATAGAGCCCACTTCTCAATAGGTCTTGCAGAAACAGGAGCA 3657
QY 3161 GACTACACAGGCGCAAGTCTCAACTTGGCAAGCTCTCCGTAGTTGAGCTTTCATGGAATTT 3220
Db 3658 GACTACACAGGCGCAAGTCTCAACTTGGCAAGCTCTCCGTAGTTGAGCTTTCATGGAATTT 3717
QY 3221 CTGAGCTTATCTATAGAGATCAATGGCTCTCTTAACGTTTCTGCAATTCGCAATTCGATAG 3280
Db 3718 CTGAGCTTATCTATAGAGATCAATGGCTCTCTTAACGTTTCTGCAATTCGCAATTCGATAG 3777
QY 3281 TTCTCTTGGGAAAGGAGCTTCTTAACAGAGGTGTCTTCAAGCGACTCCATCAAACTTGGC 3340
Db 3778 TTCTCTTGGGAAAGGAGCTTCTTAACAGAGGTGTCTTCAAGCGACTCCATCAAACTTGGC 3837
QY 3341 AGTATGTTCTAAAGGGGCTTTGTACCATTTTTGGAGGAACATTTGTGCAAAATTTGTTTAC 3400

Qy	5561	GAACTGACAAATAGTGTGCTCGTGTGTTTGGAGTCATCTTTGTATGAATAAATCTAGTCTT	5620	Qy	6641	ACGACTTATCCCACTCGCAGCAGCCACTGCTAAACAGGATTAGCAGAGCAGGATGTAG	6700
Db	6046	GAACTGACAAATAGTGTGCTCGTGTGTTTGGAGTCATCTTTGTATGAATAAATCTAGTCTT	6105	Db	7126	ACGACTTATCCCACTCGCAGCAGCCACTGCTAAACAGGATTAGCAGAGCAGGATGTAG	7185
Qy	5621	TGATCTAAATAATCTTGACGAGCCAAAGCGATAAATACCCAAATCTAAATCTTTTAA	5680	Qy	6701	CGGTGCTACAGAGTCTTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGCAGATAT	6760
Db	6106	TGATCTAAATAATCTTGACGAGCCAAAGCGATAAATACCCAAATCTAAATCTTTTAA	6165	Db	7186	CGGTGCTACAGAGTCTTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGCAGATAT	7245
Qy	5581	ACGTAAAGGACAGATGCTGCTGTATTAACCCCAATCAGCTGTAGTCTGATC	5740	Qy	6761	TTGTATCTGGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGAT	6820
Db	6166	ACGTAAAGGACAGATGCTGCTGTATTAACCCCAATCAGCTGTAGTCTGATC	6225	Db	7246	TTGTATCTGGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGAT	7305
Qy	5741	CTCATCAACTTGGGGCAGTATCTTCTTTTAGAGAAATTTGGGAGATCGGATATCGAG	5800	Qy	6821	CGGCAAAACAAACACCGCTGTAGCGGTGTTTGTTCGAAGCAGCAGATTAACGC	6880
Db	6226	CTCATCAACTTGGGGCAGTATCTTCTTTTAGAGAAATTTGGGAGATCGGATATCGAG	6285	Db	7306	CGGCAAAACAAACACCGCTGTAGCGGTGTTTGTTCGAAGCAGCAGATTAACGC	7365
Qy	5801	AAAAGGTACGCTGAATTTTAAACGTGAAATTTATCTCAAGATCTCTGCCCTCGCGCGTTT	5860	Qy	6881	GCAGAAAAAAGGATCTCAAGAGATCCTTTTGATCTTTCTACGGGGTCTGAOCCTCAGT	6940
Db	6286	AAAAGGTACGCTGAATTTTAAACGTGAAATTTATCTCAAGATCTCTGCCCTCGCGCGTTT	6345	Db	7366	GCAGAAAAAAGGATCTCAAGAGATCCTTTTGATCTTTCTACGGGGTCTGAOCCTCAGT	7425
Qy	5861	GGTGATCAGCGTGAACCTCTGACATGACAGCTCCCGGAGACGGTCAAGCTTTGCTG	5920	Qy	6941	GGAACGAAAACTCAGCTTAAGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACT	7000
Db	6346	GGTGATCAGCGTGAACCTCTGACATGACAGCTCCCGGAGACGGTCAAGCTTTGCTG	6405	Db	7426	GGAACGAAAACTCAGCTTAAGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACT	7485
Qy	5921	TAAGCGATGCGGGAGCAGAAAGCCGTCAGGGCGCGTCAGGGGTGTTGGCGGTGT	5980	Qy	7001	AGATCCTTTT	7010
Db	6406	TAAGCGATGCGGGAGCAGAAAGCCGTCAGGGCGCGTCAGGGGTGTTGGCGGTGT	6465	Db	7486	AGATCCTTTTACGGCGCCCTGTAGCGGCCATTAAGCGCGGGGTGTGGTGTACCGGC	7545
Qy	5981	CGGGCGCAGCATGACCGTACGTACGTAGCGAGTGTATCTGCTTAACTATG	6040	Qy	7011	-----	7010
Db	6466	CGGGCGCAGCATGACCGTACGTACGTAGCGAGTGTATCTGCTTAACTATG	6525	Db	7546	AGCGTGACCGCTACACTTGCACGCGCCCTAGCGCCCGCTCTTTCGCTTTCTTCCCTTCC	7605
Qy	6041	CGGCATCAGACAGATTGTACTGAGAGTGCAACCATATGCGGTGTGAATACCGCACAGAT	6100	Qy	7011	-----	7010
Db	6526	CGGCATCAGACAGATTGTACTGAGAGTGCAACCATATGCGGTGTGAATACCGCACAGAT	6585	Db	7606	TTTTCGCCACAGTTCGCCGGCTTTCCCGCTCAAGCTCTAAATCGGGGCTCCCTTTAGGG	7665
Qy	6101	CGTAAGGAGAAATACCGCATCAGGCGCTTCTCCGCTTCTCGCTCACTGACTCGCTGC	6160	Qy	7011	-----	7010
Db	6586	CGTAAGGAGAAATACCGCATCAGGCGCTTCTCCGCTTCTCGCTCACTGACTCGCTGC	6645	Db	7666	TTTTCGCCACAGTTCGCCGGCTTTCCCGCTCAAGCTCTAAATCGGGGCTCCCTTTAGGG	7725
Qy	6161	GCTCGCTGCTTCCGCTGCGCGAGCGGTATCAGCTCACTCAAAGCGGTAAATACGGTTAT	6220	Qy	7011	-----	7010
Db	6646	GCTCGCTGCTTCCGCTGCGCGAGCGGTATCAGCTCACTCAAAGCGGTAAATACGGTTAT	6705	Db	7726	CGTAGTGGCCATCGCCCTGTATGACGGTTCCTCCCGCTTTGACGTTGGAGTCCACGCTC	7785
Qy	6221	CCACGAATCAGGGATTAACCGCAGGAAGAACATGTGAGCAAAAGGCGCAGCAAGGCCA	6280	Qy	7011	-----	7010
Db	6706	CCACGAATCAGGGATTAACCGCAGGAAGAACATGTGAGCAAAAGGCGCAGCAAGGCCA	6765	Db	7786	TTTAAATAGTGGACTCTTGTTCAAACTCGAACAACACTCAACCTATCTCGGTCTATTCT	7845
Qy	6281	GGAACGTTAAAGCCGCTTGTGCGGTTCATAGGCTCCGCCCGCCCTGACGAGC	6340	Qy	7011	-----	7010
Db	6766	GGAACGTTAAAGCCGCTTGTGCGGTTCATAGGCTCCGCCCGCCCTGACGAGC	6825	Db	7846	TTTGAATTAAGGGATTTTCGGATTTTCGGCTTAAAGATGAGCTGATTTAA	7905
Qy	6341	ATCAAAAAATCAGCTCAAGTCAGAGTGGGAAACCGGACAGGACTATAAGATACC	6400	Qy	7011	-----	7010
Db	6826	ATCAAAAAATCAGCTCAAGTCAGAGTGGGAAACCGGACAGGACTATAAGATACC	6885	Db	7906	CAAAAATTAACCGCAATTTTAAACAAATATTAACGTTTACATTTAAATCAATCTAAAG	7965
Qy	6401	AGGGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTTTCGACACCTGCGGCTTACCG	6460	Qy	7044	TATATATAGTAAATCTTGGTGTGACAGTTACCAATGCTTAATCAGTAGGACCTATCTC	7103
Db	6886	AGGGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTTTCGACACCTGCGGCTTACCG	6945	Db	7966	TATATATAGTAAATCTTGGTGTGACAGTTACCAATGCTTAATCAGTAGGACCTATCTC	8025
Qy	6461	GATACCTGCGGCTTTCTCCCTTGGGAAGCGTGGCGCTTTCTATAGCTCAGCTGTA	6520	Qy	7104	ACGGATCTGTCTATTTCGTTTCATCCATAGTTGCTGACTCCCGCTGCTGTAGATAACTAC	7163
Db	6946	GATACCTGCGGCTTTCTCCCTTGGGAAGCGTGGCGCTTTCTATAGCTCAGCTGTA	7005	Db	8026	ACGGATCTGTCTATTTCGTTTCATCCATAGTTGCTGACTCCCGCTGCTGTAGATAACTAC	8085
Qy	6521	GGTATCTCAGTTCGGTGTAGTGTGCTCCAGCTGGGCTGTGACAGAACCCCGG	6580	Qy	7164	GATACGGGAGGCTTACCATCTGGCCCCAGTGTCTCAATGATACCGGAGACCCACGCTC	7223
Db	7006	GGTATCTCAGTTCGGTGTAGTGTGCTCCAGCTGGGCTGTGACAGAACCCCGG	7065	Db	8086	GATACGGGAGGCTTACCATCTGGCCCCAGTGTCTCAATGATACCGGAGACCCACGCTC	8145
Qy	6581	TTACGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTTAGTCAACCCCGTAAAGC	6640	Qy	7224	ACGGCTCCAGATTTATCAGCAATAAACCCAGCAGCGGAAAGGCGGAGCCAGAGTGG	7283
Db	7066	TTACGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTTAGTCAACCCCGTAAAGC	7125	Db	8146	ACGGCTCCAGATTTATCAGCAATAAACCCAGCAGCGGAAAGGCGGAGCCAGAGTGG	8205
				Qy	7284	TCTGCAACTTTATCCGCTCCATCCAGTCTATTAAATTTGTTCCGGGAAGCTAGAGTAAG	7343

RESULT 7
AAD05811
ID AAD0
XX
AC AAD0

QY	1561	ATCGAGCGCTTCATTTTGTGATACCTTTTATTTGTAAACCTATATAGTATAGGATTTTTTTT	1620
DB	1355	ATCGAGCGCTTCATTTTGTACTTTTTTATTTGTAACTTATATAGTATAGGATTTTTTTT	1414
QY	1621	GTCAATTTGTCTTCTCGTACGAGCTTGCTCCTGATCAGCCCTATCTCCAGCTGATGAA	1680
DB	1415	GTCAATTTGTCTTCTCGTACGAGCTTGCTCCTGATCAGCCCTATCTCCAGCTGATGAA	1474
QY	1681	TATCTTGTGTGATGGGTTTCGGGAAATCAATCGAGTTGATGTTTTCTTGGTATTTCC	1740
DB	1475	TATCTTGTGTGATGGGTTTCGGGAAATCAATCGAGTTGATGTTTTCTTGGTATTTCC	1534
QY	1741	ACTCCCTCTTCAGAGTACAGAAAGATTAAGTGTGAGAAAGTTCTGTTGTGCAAGCTTATCGATAA	1800
DB	1535	ACTCCCTCTTCAGAGTACAGAAAGATTAAGTGTGAGAAAGTTCTGTTGTGCAAGCTTATCGATAA	1594
QY	1801	GCTTTAATCGGTAGTTTTATCA CAGTTAAATTCCTAAACGGAGTCAAGGCAACCGTGATGAA	1860
DB	1595	GCTTTAATCGGTAGTTTTATCA CAGTTAAATTCCTAAACGGAGTCAAGGCAACCGTGATGAA	1654
QY	1861	ATCTAACAATGCGCTCATCGTCTCATCTCGGCACCGTCAACCTCGGATGCTGTAGGCATAG	1920
DB	1655	ATCTAACAATGCGCTCATCGTCTCATCTCGGCACCGTCAACCTCGGATGCTGTAGGCATAG	1714
QY	1921	CTTGGTTATGCCGGTACTGCGGGGCCCTTTGCGGGATATGTCCTTCCAGACAGCATCGC	1980
DB	1715	CTTGGTTATGCCGGTACTGCGGGGCCCTTTGCGGGATATGTCCTTCCAGACAGCATCGC	1774
QY	1981	CAGTCACTATGGCGTGCTGCTAGCGCTATATGGTTGTATGTCAAATTTCTATGCGCACCCGT	2040
DB	1775	CAGTCACTATGGCGTGCTGCTAGCGCTATATGGTTGTATGTCAAATTTCTATGCGCACCCGT	1834
QY	2041	TCTCGGAGCACTGTCCGACCGCTTTGGCGCGCCGCCAGTCTGCTCGCTTCGCTACTTGG	2100
DB	1835	TCTCGGAGCACTGTCCGACCGCTTTGGCGCGCCGCCAGTCTGCTCGCTTCGCTACTTGG	1894
QY	2101	AGCCACTATCGACTACGCGATCATGGCGGACCAACACCGCTCCTGTGATCTATCGATCTA	2160
DB	1895	AGCCACTATCGACTACGCGATCATGGCGGACCAACACCGCTCCTGTGATCTATCGATCTA	1954
QY	2161	AATGTAAAGTTAAATCTCTAAATAATTAATAAGTCCAGTTTCTCCATACGAACCTTAA	2220
DB	1955	AATGTAAAGTTAAATCTCTAAATAATTAATAAGTCCAGTTTCTCCATACGAACCTTAA	2014
QY	2221	CAGCATTTGCGGTGAGCATCTAGACCTTCAACAGGAGCCAGATCCATCACTGCTTGGCCAA	2280
DB	2015	CAGCATTTGCGGTGAGCATCTAGACCTTCAACAGGAGCCAGATCCATCACTGCTTGGCCAA	2074
QY	2281	TATGTTTTCACTCCCTCAGGAGTTACGTTCTTGTGAAGTGTAAAGTCTCTGGAAGTTTCAG	2340
DB	2075	TATGTTTTCACTCCCTCAGGAGTTACGTTCTTGTGAAGTGTAAAGTCTCTGGAAGTTTCAG	2134
QY	2341	TGTTAACTCCGCTGTATGTAGCGGCGATATCGGTACGTTGCGCAAGTGTGTTTGGTACCGG	2400
DB	2135	TGTTAACTCCGCTGTATGTAGCGGCGATATCGGTACGTTGCGCAAGTGTGTTTGGTACCGG	2194
QY	2401	AGGAGTAACTCCCAAACTCTCTGGAGGTAGGCACCAACAAACAGATCCAGCGTGT	2460
DB	2195	AGGAGTAACTCTCCAACTCTCTGGAGGTAGGCACCAACAAACAGATCCAGCGTGT	2254
QY	2461	GTACTTGATCAACATAAGAAAGCATTTCTCGATTTTCAGGATCAAGTGTTCAGGAGCGT	2520
DB	2255	GTACTTGATCAACATAAGAAAGCATTTCTCGATTTTCAGGATCAAGTGTTCAGGAGCGT	2314
QY	2521	ACTGATTGGACATTTCCAAAGCCCTGCTCGTAGGTTGCAACCGATAGGGTTGTAGAGTGTG	2580
DB	2315	ACTGATTGGACATTTCCAAAGCCCTGCTCGTAGGTTGCAACCGATAGGGTTGTAGAGTGTG	2374
QY	2581	CAATACACTTGGTACAAATTTCAACCCCTTGGCAACTCGACAGCTTGGTTGTGAACAGCAT	2640
DB	2375	CAATACACTTGGTACAAATTTCAACCCCTTGGCAACTCGACAGCTTGGTTGTGAACAGCAT	2434

2641	QY	CTTCAATTCCTGGCAAGCTCCTTGTCTGTCTCATATCGACGCCAACAGAAATCACCTGGGAAT	2700
2435	DB	CTTCAATTCCTGGCAAGCTCCTTGTCTGTCTCATATCGACGCCAACAGAAATCACCTGGGAAT	2494
2701	QY	CAATACCATGTTTCTCAGCTTTGAGACAGAAGGTCTTGAGCCAAACGAAATCTGGATCAGCGTATT	2760
2495	DB	CAATACCATGTTTCTCAGCTTTGAGACAGAAGGTCTTGAGCCAAACGAAATCTGGATCAGCGTATT	2554
2761	QY	TATCAGCAATAACTAGAACTTTCAGAAAGGCCACGACGGCATGTCAATTAATCTACACAGGGCTG	2820
2555	DB	TATCAGCAATAACTAGAACTTTCAGAAAGGCCACGACGGCATGTCAATTAATCTACACAGGGCTG	2614
2821	QY	ATGTGTCAATTTTGAAACCATCACTCTTGCGCAGCAGTAACGAACTGGTTTCTCTGGACCAATA	2880
2615	DB	ATGTGTCAATTTTGAAACCATCACTCTTGCGCAGCAGTAACGAACTGGTTTCTCTGGACCAATA	2674
2881	QY	TTTTGTCACTTTAGGAAACAGTTTCTTGTTCGTAAAGCCATAGACAGCTACTGCGCTGGGCGC	2940
2675	DB	TTTTGTCACTTTAGGAAACAGTTTCTTGTTCGTTCGTAAAGCCATAGACAGCTACTGCGCTGGGCGC	2734
2941	QY	CTCCTCGTAGCAGTACACTTAGCACCACCTTTGCGGCAACGCTAGATGACTTCTGCGG	3000
2735	DB	CTCCTCGTAGCAGTACACTTAGCACCACCTTTGCGGCAACGCTAGATGACTTCTGCGG	2794
3001	QY	TAAAGGTACCATCCTTCTTAGTGGAGATGCAAAAAACAATTTCTTTTGCACACGACAACTT	3060
2795	DB	TAAAGGTACCATCCTTCTTAGTGGAGATGCAAAAAACAATTTCTTTTGCACACGACAACTT	2854
3061	QY	TGCGAGAACACCCACGATCAGGGAAGTGGAGGACAGAAATGCGGTTCACCCACGAGGAAT	3120
2855	DB	TGCGAGGAACACCCACGATCAGGGAAGTGGAGGACAGAAATGCGGTTCACCCACGAGGAAT	2914
3121	QY	AGAGGCCAACTTCTCAATAGGTCTTGCAAAAACGAGACAGACTACACACGAGGCAAGTCT	3180
2915	DB	AGAGGCCAACTTCTCAATAGGTCTTGCAAAAACGAGACAGACTACACACGAGGCAAGTCT	2974
3181	QY	CAACTTGCACAGCTCTCGTTAGTTGAGCTTATGGAATTTCTTCTGACGTTATCTATAGAGA	3240
2975	DB	CAACTTGCACAGCTCTCGTTAGTTGAGCTTATGGAATTTCTTCTGACGTTATCTATAGAGA	3034
3241	QY	GATCAATGGCTCTCTTAAACGTTATCTGGCAATTTGCATAGTTCCTCTGGGAAAGGAGCTT	3300
3035	DB	GATCAATGGCTCTCTTAAACGTTATCTGGCAATTTGCATAGTTCCTCTGGGAAAGGAGCTT	3094
3301	QY	CTAACACAGGTGCTTCAAAGCGACTCCATCAAACTTGGCAGTTAGTTCTTAAAGGGCTT	3360
3095	DB	CTAACACAGGTGCTTCAAAGCGACTCCATCAAACTTGGCAGTTAGTTCTTAAAGGGCTT	3154
3361	QY	TGTCACCAATTTGACGACAACTTTCGACAAATTTGGTTGACTAATTCATATCTGTTCCG	3420
3155	DB	TGTCACCAATTTGACGACAACTTTCGACAAATTTGGTTGACTAATTCATATCTGTTCCG	3214
3421	QY	TTTTTCTGGATAGGACACGAAGGCGATCTTCAATTTCTTCTGAGGAGCCCTTAGAAACGT	3480
3215	DB	TTTTTCTGGATAGGACACGAAGGCGATCTTCAATTTCTTCTGAGGAGCCCTTAGAAACGT	3274
3481	QY	CAATTTTGCACAAATCAATACGACCTTTCAGAGGACCTTCTTTAGTTTGGATTCCTTCTT	3540
3275	DB	CAATTTTGCACAAATCAATACGACCTTTCAGAGGACCTTCTTTAGTTTGGATTCCTTCTT	3334
3541	QY	TAGTGTTCCTTGGGTATCTCTGGCTTGGCATCTCCTTTCTTCTAGTCACCTTTAGGG	3600
3335	DB	TAGTGTTCCTTGGGTATCTCTGGCTTGGCATCTCCTTTCTTCTAGTCACCTTTAGGG	3394
3601	QY	ACTTTCATATCCAGGTTTCTCTCCACCTCGTCCAAACGTCACACCGTACTTGGCACATCTAA	3660
3395	DB	ACTTTCATATCCAGGTTTCTCTCCACCTCGTCCAAACGTCACACCGTACTTGGCACATCTAA	3454
3661	QY	CTAATGCAAAATAAATAAGTCAGCACATTTCCAGGCTATATCTTCTCTGGATTAGCTT	3720
3455	DB	CTAATGCAAAATAAATAAGTCAGCACATTTCCAGGCTATATCTTCTCTGGATTAGCTT	3514
3721	QY	CTGCAAGTTCATCAGCTTCTCTCCTAATTTTAGCGTTCAACAAAACCTCGTCGTCAAATA	3780

Db 3515 CTGCAAGTTCTATCAGCTTCTCCCTAATTTTAGGTTCAACAAAACCTTCGTCTCAANTA 3574
QY
Db 3781 ACCGTTGGTATAGAACCTTCTGGAGCATGCTCTTACGATCCCAAGAGTGGCTTCCA 3840
Db 3575 ACCGTTGGTATAGAACCTTCTGGAGCATGCTCTTACGATCCCAAGAGTGGCTTCCA 3634
QY 3841 TGGCTCTAAGACCTTTGATTTGGCCAAACAGGAAGTGGCTTCAAGTGCACAAACCAA 3900
Db 3635 TGGCTCTAAGACCTTTGATTTGGCCAAACAGGAAGTGGCTTCAAGTGCACAAACCAA 3694
QY 3901 CACCTGTTGTTCAACACAAAATTTCAAGCAGTCTCCATCACAATCCAAATTCGATACCCA 3960
Db 3695 CACCTGTTGTTCAACACAAAATTTCAAGCAGTCTCCATCACAATCCAAATTCGATACCCA 3754
QY 3961 GCAACTTTTGAGTTGCTCCAGATGACACCTTTTATACCAACCAACCGTGCAGCAGATT 4020
Db 3755 GCAACTTTTGAGTTGCTCCAGATGACACCTTTTATACCAACCAACCGTGCAGCAGATT 3814
QY 4021 GGTAGACTCCAGTTTGTGCTTATAGCTCCGGAATAGACTTTTGGACGAGTACACCA 4080
Db 3815 GGTAGACTCCAGTTTGTGCTTATAGCTCCGGAATAGACTTTTGGACGAGTACACCA 3874
QY 4081 GGCCCAACGAGTAATTAGAGAGTGCAGCACCACAAAGTAGTGAATAGACCAATCGGGGGGT 4140
Db 3875 GGCCCAACGAGTAATTAGAGAGTGCAGCACCACAAAGTAGTGAATAGACCAATCGGGGGGT 3934
QY 4141 CAGTAGTCAAGAGCCCAACAAAATTTCACTGACAGGAACTTTTGACATCTTCAGAAA 4200
Db 3935 CAGTAGTCAAGAGCCCAACAAAATTTCACTGACAGGAACTTTTGACATCTTCAGAAA 3994
QY 4201 GTTCGTAATTCAGTAGTCAATTCGCGAGATCAATAATGGGGAATTATACCAAGAGCAACAG 4260
Db 3995 GTTCGTAATTCAGTAGTCAATTCGCGAGATCAATAATGGGGAATTATACCAAGAGCAACAG 4054
QY 4261 TGGAGTCAACATCTACCAACTTTTCGGTCTCAGAAAAGCAATAGTCTTACTACCGC 4320
Db 4055 TGGAGTCAACATCTACCAACTTTTCGGTCTCAGAAAAGCAATAGTCTTACTACCGC 4114
QY 4321 CATTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAAGGCACAGCCATACTAGCAT 4380
Db 4115 CATTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAAGGCACAGCCATACTAGCAT 4174
QY 4381 TAGGGGCAAGGATGCAACTTTTATCAACAGGCTCTATAGATTAACCTAGCCCTGGGA 4440
Db 4175 TAGGGGCAAGGATGCAACTTTTATCAACAGGCTCTATAGATTAACCTAGCCCTGGGA 4234
QY 4441 TCATCTTTTGGACAACTCTTCTGCGCAATCTAGGTCACAAATCACTTCATTGATACCAT 4500
Db 4235 TCATCTTTTGGACAACTCTTCTGCGCAATCTAGGTCACAAATCACTTCATTGATACCAT 4294
QY 4501 TATTGTACAACTTGAGCAAGTTGCGATCAGCTCTCAAAATGGTCTCTGTAAACGATG 4560
Db 4295 TATTGTACAACTTGAGCAAGTTGCGATCAGCTCTCAAAATGGTCTCTGTAAACGATG 4354
QY 4561 ACTCAACTTTGACATTAATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 4620
Db 4355 ACTCAACTTTGACATTAATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 4414
QY 4621 GCTGGTACAGCAGTATAGGAAACACGCTTTTCTACCAAACTCAAGAAATTTATCAAACT 4680
Db 4415 GCTGGTACAGCAGTATAGGAAACACGCTTTTCTACCAAACTCAAGAAATTTATCAAACT 4474
QY 4681 CTGCAACATTTGCGTATGCGAGTACGAGGAATGTCATACTTGAAGTGGACAGTGAG 4740
Db 4475 CTGCAACATTTGCGTATGCGAGTACGAGGAATGTCATACTTGAAGTGGACAGTGAG 4534
QY 4741 TGTAGTCTTGAGAAATTTCTGAAGCGGATTTTATTATCAAGTCAAGTCAAGTCAAGTCAAGTCA 4800
Db 4535 TGTAGTCTTGAGAAATTTCTGAAGCGGATTTTATTATCAAGTCAAGTCAAGTCAAGTCAAGTCA 4594
QY 4801 TCCTCTACCGGAGGATCGTGGCGA----- 4828

Db 4595 TCCTCTACCGGAGCGCATCGTGGCCGACCTGCGCGAGCTGCAGGGGGGGGGGGCGCTGAGTCTGC 4654
QY 4829 ----- 4828
Db 4655 CTCGTGAAGAAGAGTGTCTGCTGACTATACAGGCTGAAATCGCCCCATCATCCAGCCAGA 4714
QY 4829 ----- 4828
Db 4715 AAGTCAGGAGCCACGGTTGATGAGAGCTTTGTTGTAGGTGGAACCAAGTGTGATTTTGA 4774
QY 4829 ----- 4828
Db 4775 ACTTTTGTCTTCCACGGAACGGTCTGCGTTGTCGGGAAGATGCGTGAATCTGATCTCTCA 4834
QY 4829 ----- 4828
Db 4835 ACTCAGCAAAAGCTTCGATTTATTCAACAAAGCCGCTCCGTCAGAGTCAGGTAATGCT 4894
QY 4829 ----- 4828
Db 4895 CTGCCAGTGTACACCAATTAACCAATTTCTGATTAGAAAAAATCATCGAGCATCAAAATG 4954
QY 4829 ----- 4828
Db 4955 AAATGCAATTTATTCATATCAGGATTAATCAATACCATATTTTGAATAAAGCCGTTCTG 5014
QY 4829 ----- 4828
Db 5015 TAATGAAGGAGAAAACTCACCGAGCAGTCTCCATAGGATGGCAAGATCTCGTATCGGTC 5074
QY 4829 ----- 4828
Db 5075 TCGGATTCGGACTCGTCCAAACATCAATACACCTATTAAATTTCCCTTCGTCAAAAATAAG 5134
QY 4829 ----- 4828
Db 5135 GTTATCAAGTGAGAAATCACCATGATGACGACTGAATCCGCTGAGAAATGGCAAAAGCTT 5194
QY 4829 ----- 4828
Db 5195 ATGCATTTCTTCCAGACTTGTTCACAGGCGCCATACGCTGTCATCAAAATCACT 5254
QY 4829 ----- 4828
Db 5255 CGCATCAACCAACCGTTATTTCATTTCGTTGATTGCGCTGAGCGAGAGCAAAATACGCGATC 5314
QY 4829 ----- 4828
Db 5315 GCTGTTAAAGGACAATTAACAAAGGAATCGAATGCAACCGCGCGAGGAACACTGCCAG 5374
QY 4829 ----- 4828
Db 5375 CCATCAACAAATATTTTCACTGATCAGGATATTTCTTAATACCTGGAATGCTGTTTT 5434
QY 4829 ----- 4828
Db 5435 CCGGGGATCGCAGTGGTGAATAACCATGTCATCAGGATACGGATAAAATGCTTGAT 5494
QY 4829 ----- 4828
Db 5495 GGTGGAAGAGGCATAAAATTCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAAACATC 5554
QY 4829 ----- 4828
Db 5555 ATTGGCAACGCTACCTTTTCCCATGTTTCAGAAAACAATCTCGCGCATCGGCTTCCCATATA 5614
QY 4829 ----- 4828
Db 5615 CAATCGATAGATTGTCGACCTGATTGCCGACATTAATCGCGAGCCCATTTTATACCCATA 5674
QY 4829 ----- 4828
Db 5675 TAAATCAGCATCCATGTTTGAATTTAATCGCGGCTCGAGCAAGACGTTTCCCGTTGAAT 5734

Db 7895 TTTTCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCTCTTTGATCTTTT 7954
Qy 6920 CTACGGGGTCTGAGCCTCAGTGGAAAGCAAACTCAGCTTAAGGATTTTGGTTCATGAGAT 6979
Db 7955 CTACGGGGTCTGAGCCTCAGTGGAAAGCAAACTCAGCTTAAGGATTTTGGTTCATGAGAT 8014
Qy 6980 TATCAAAAAAGGATCTTCACTAGATCTCTTTAAATTAATAAATGAAGTTTAAATCAATCT 7039
Db 8015 TATCAAAAAAGGATCTTCACTAGATCTCTTTAAATTAATAAATGAAGTTTAAATCAATCT 8074
Qy 7040 AAGTATATATAGTAATCTGCTGACAGTGTACCAATCTTAATCAAGTGAGGACCTTA 7099
Db 8075 AAGTATATATAGTAATCTGCTGACAGTGTACCAATCTTAATCAAGTGAGGACCTTA 8134
Qy 7100 TCTCAGGATCTGCTCTATTTGGTTTCACTCATAGTTGCTGACTCCCGTCTGTAGATAA 7159
Db 8135 TCTCAGGATCTGCTCTATTTGGTTTCACTCATAGTTGCTGACTCCCGTCTGTAGATAA 8194
Qy 7160 CTACGATACGGGAGGCTTACCATCTGCGCCAGTGTCTGAATGATACCGGAGCCAC 7219
Db 8195 CTACGATACGGGAGGCTTACCATCTGCGCCAGTGTCTGAATGATACCGGAGCCAC 8254
Qy 7220 GCTCACCGGCTCCAGATTATACCAATTAACACGAGCGGAGGCGGAGCGAGAA 7279
Db 8255 GCTCACCGGCTCCAGATTATACCAATTAACACGAGCGGAGGCGGAGCGAGAA 8314
Qy 7280 GTGGTCTTGCACCTTATCCGCTCCATCCAGTCTTAATTTGTTGCGGAGAGCTAGAG 7339
Db 8315 GTGGTCTTGCACCTTATCCGCTCCATCCAGTCTTAATTTGTTGCGGAGAGCTAGAG 8374
Qy 7340 TAAGTATGTCGAGTAAATAGTTTGGCAACGTTTGCATTGCTGAGGAGCATCTGG 7399
Db 8375 TAAGTATGTCGAGTAAATAGTTTGGCAACGTTTGCATTGCTGAGGAGCATCTGG 8434
Qy 7400 TGTCAAGCTCGTCTGTTGGTATGCTTCATTCACTCGGTTCCCAACGATCAAGGCGAG 7459
Db 8435 TGTCAAGCTCGTCTGTTGGTATGCTTCATTCACTCGGTTCCCAACGATCAAGGCGAG 8494
Qy 7460 TTACATGATCCGCTATGTTGCAAAAGGGTGTAGTCTCTCGTCTCCGATCTTG 7519
Db 8495 TTACATGATCCGCTATGTTGCAAAAGGGTGTAGTCTCTCGTCTCCGATCTTG 8554
Qy 7520 TCAGAAATAGTGGCCGAGTGTATCACTCATGCTGTTATGGCAGCACTGCATAATCTC 7579
Db 8555 TCAGAAATAGTGGCCGAGTGTATCACTCATGCTGTTATGGCAGCACTGCATAATCTC 8614
Qy 7580 TTACTGTATGCTCATGCTGCTTCTGTGATCTGTGAGTACTCAACCAAGTCAT 7639
Db 8615 TTACTGTATGCTCATGCTGCTTCTGTGATCTGTGAGTACTCAACCAAGTCAT 8674
Qy 7640 TCTGAGATAGTGTATGCGGAGCGAGTGTCTCTTGGCCGCGCTCAACACGGATTAATA 7699
Db 8675 TCTGAGATAGTGTATGCGGAGCGAGTGTCTCTTGGCCGCGCTCAACACGGATTAATA 8734
Qy 7700 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAAAAGTCTTTCGGGGCGAA 7759
Db 8735 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAAAAGTCTTTCGGGGCGAA 8794
Qy 7760 AACTCTCAGGATCTTACGCTGTTGAGATCCAGTGTGATCCCACTGCTCACCCA 7819
Db 8795 AACTCTCAGGATCTTACGCTGTTGAGATCCAGTGTGATCCCACTGCTCACCCA 8854
Qy 7820 ACTGATCTTCAAGTCTTTTACTTTTACCAGGTTTCTGGGTGAGCAAAAAACAGGAGGC 7879
Db 8855 ACTGATCTTCAAGTCTTTTACTTTTACCAGGTTTCTGGGTGAGCAAAAAACAGGAGGC 8914
Qy 7880 AAAATGCCGCAAAAAGGATTAAGGCGACACCGAAATGTTGATCTACTCTCTTCC 7939
Db 8915 AAAATGCCGCAAAAAGGATTAAGGCGACACCGAAATGTTGATCTACTCTCTTCC 8974
Qy 7940 TTTTTCATATTAATGAAGCATTTATCAGGGTTATTTGCTCATGAGCGGATACATATTG 7999

Db 8975 TTTTTCATATTAATGAAGCATTTATCAGGGTTATTTGCTCATGAGCGGATACATATTG 9034
Qy 8000 AATGTTATTAAGAAAAATAAATAAGGTTTCGCGCACATTTCCCGAAAAAGTGGCAC 8059
Db 9035 AATGTTATTAAGAAAAATAAATAAGGTTTCGCGCACATTTCCCGAAAAAGTGGCAC 9094
Qy 8060 CTGAGCTCTAAGAACCAATTTATCATGATTAACCTATATAAAATAGCGTATCAGA 8119
Db 9095 CTGAGCTCTAAGAACCAATTTATCATGATTAACCTATATAAAATAGCGTATCAGA 9154
Qy 8120 GGGCTTTTCTTCAAGAAATTAATTTCTATGTTGACAGCTTATCATCGAATAGCTGAC 8179
Db 9155 GGGCTTTTCTTCAAGAAATTAATTTCTATGTTGACAGCTTATCATCGAATAGCTGAC 9214
Qy 8180 TCATGTTGTTTGTGAATAGACGAGATCGGAAACACTGAAAAATAACAGTTTATTATT 8239
Db 9215 TCATGTTGTTTGTGAATAGACGAGATCGGAAACACTGAAAAATAACAGTTTATTATT 9274
Qy 8240 CG 8241
Db 9275 CG 9276
RESULT 8
AAD13151
ID AAD13151 standard; DNA; 9276 BP.
XX
AC AAD13151;
XX
DT 16-OCT-2001 (first entry)
XX
DE Pichia pastoris pPIC9K vector DNA.
XX
KW Human; transmembrane serine protease; membrane-type serine protease;
KW MTSP; protease domain; neoplastic disease; tumour; cancer; cytostatic;
KW lung carcinoma; colon adenocarcinoma; ovarian carcinoma; gene therapy;
KW cyclic; circular; ds.
XX
OS Pichia pastoris.
XX
PN WO200157194-A2.
XX
XX 09-AUG-2001.
FD
XX 02-FEB-2001; 2001WO-US003471.
XX
XX 03-FEB-2000; 2000US-0179982P.
PR 18-FEB-2000; 2000US-0183542P.
PR 22-JUN-2000; 2000US-0213124P.
PR 26-JUL-2000; 2000US-0220970P.
PR 08-SEP-2000; 2000US-00657986.
PR 22-SEP-2000; 2000US-0234840P.
XX
XX (CORV-) CORVAS INT INC.
XX
PI Madison EL, Ong EO, Yeh J;
XX
XX MPI; 2001-488877/53.
XX
PT Novel single chain polypeptide comprising protease domain of type-II
PT membrane-type serine protease or its catalytically active portion useful
PT for treating and preventing cancer and tumor.
XX
PS Example 4; Page 221-223; 256pp; English.
XX
CC The invention relates to transmembrane serine proteases and their
CC corresponding nucleotides and the protease domain of a type-II membrane-
CC type serine protease (MTSP). MTSP is useful for identifying compounds
CC that modulate or inhibit its proteolytic activity and for formulating a
CC medicament for treating neoplastic disease. MTSP and its corresponding
CC nucleotides are useful in preventing or treating tumours or cancers such
CC as lung carcinoma, colon adenocarcinoma and ovarian carcinoma, in
CC diagnostics and in hybridisation assays. MTSP is useful as a diagnostic

CC marker for tumour development, growth and/or progression and as
CC immunogens to generate antibodies that specifically bind to it. MTSP DNA
CC is useful in a yeast two-hybrid system and in gene therapy. The present
CC sequence is Pichia pastoris pPIC9K vector DNA
xx
SQ Sequence 9276 BP; 2493 A; 2197 C; 2078 G; 2508 T; 0 U; 0 Other;

```
Query Match          79.6%; Score 6561.6; DB 4; Length 9276;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;

QY 1 AGATCTAACATCCAAAGAGCAAGAGGTTGAATGAAACCTTTTGGCATCCGACATCCACAG 60
DB 1 AGATCTAACATCCAAAGAGCAAGAGGTTGAATGAAACCTTTTGGCATCCGACATCCACAG 60
QY 61 GTCCATTCTCACATAGTSCCAAAACGCAACAGAGGGGATACACTAGCAGCAGACCGT 120
DB 61 GTCCATTCTCACATAGTSCCAAAACGCAACAGAGGGGATACACTAGCAGCAGACCGT 120
QY 121 TGCAAACGAGAGCTCCACCTCTCTCTCTCAACACCCACTTTTGGCATCCGAAACAC 180
DB 121 TGCAAACGAGAGCTCCACCTCTCTCTCTCAACACCCACTTTTGGCATCCGAAACAC 180
QY 181 AGCCACAGTTATGGGCTTGATGGAGCTCGCTCAATTCCTTCTATTAGGCTACTA 240
DB 181 AGCCACAGTTATGGGCTTGATGGAGCTCGCTCAATTCCTTCTATTAGGCTACTA 240
QY 241 ACACATGACTTTATAGCTGTCTATCTGGGCCCCCTGGCGAGGTTCAATGTTGTTA 300
DB 241 ACACATGACTTTATAGCTGTCTATCTGGGCCCCCTGGCGAGGTTCAATGTTGTTA 300
QY 301 TTTCGGAATCAACAAAGCTCGCATACACCGCAACATCACTCCAGATGAGGGCTTCTG 360
DB 301 TTTCGGAATGCAACAAAGCTCGCATACACCGCAACATCACTCCAGATGAGGGCTTCTG 360
QY 361 AGTGCGGGTCAAAATAGTTTCATGTTCCCAAAATGGCCAAAACCTGACAGTTAAACGCT 420
DB 361 AGTGCGGGTCAAAATAGTTTCATGTTCCCAAAATGGCCAAAACCTGACAGTTAAACGCT 420
QY 421 GTCCTGGAACCTAATATGACAAAGCGTGATCTCATCCAGATCAACTAGAGTTGGTTCG 480
DB 421 GTCCTGGAACCTAATATGACAAAGCGTGATCTCATCCAGATCAACTAGAGTTGGTTCG 480
QY 481 TTGAAATGCTAACGGCCAGTTGGTCAAAAGAAACTTTCCAAAGTCGCCATACCGTTGT 540
DB 481 TTGAAATGCTAACGGCCAGTTGGTCAAAAGAAACTTTCCAAAGTCGCCATACCGTTGT 540
QY 541 CTGTTGGTATGATGACGAAATGCTCAAAATTAATCTCAATTAATGCTTAGCGAGTCT 600
DB 541 CTGTTGGTATGATGACGAAATGCTCAAAATTAATCTCAATTAATGCTTAGCGAGTCT 600
QY 601 CTCATCGCTTCTGAACCCCGTGCACTGTGCCGAAACGCAATGGGAAACACCCGCT 660
DB 601 CTCATCGCTTCTGAACCCCGTGCACTGTGCCGAAACGCAATGGGAAACACCCGCT 660
QY 661 TTTTGGATGATATGATGATGCTCCAAATGATGCTTCCAAAGATCTGGTGGGAATACT 720
DB 661 TTTTGGATGATATGATGATGCTCCAAATGATGCTTCCAAAGATCTGGTGGGAATACT 720
QY 721 GCTGATAGCTTAACGCTTCATGATCAAAATTTAACTGTTCTAACCCCTACTGTGACAGCAAT 780
DB 721 GCTGATAGCTTAACGCTTCATGATCAAAATTTAACTGTTCTAACCCCTACTGTGACAGCAAT 780
QY 781 ATATAAACAGAGGAAGCTGCCCTCTCTTAAACCTTTTTTTTTTATCATCATATTAGCTT 840
DB 781 ATATAAACAGAGGAAGCTGCCCTCTCTTAAACCTTTTTTTTTTATCATCATATTAGCTT 840
QY 841 ACTTTCATTAATGGCACTGGTTCGAATGACAGCTTTTGATTTTAACGACTTTTAACGA 900
DB 841 ACTTTCATTAATGGCACTGGTTCGAATGACAGCTTTTGATTTTAACGACTTTTAACGA 900
QY 901 CAACCTTGAGAGATCAAAAACCAACTAATTTTCGAGGATCCAAACGATCAGATTTTCT 960
DB 901 CAACCTTGAGAGATCAAAAACCAACTAATTTTCGAGGATCCAAACGATCAGATTTTCT 960
```

```
DB 901 CAACTTGAGAGATCAAAAACCAACTAATTTTCGAGGATCCAAACGATGAGATTTCT 960
QY 961 TCAATTTTACTGCACTTTTATTCGACGATCCTCCGATTAGCTGCTCCAGTCAACACT 1020
DB 961 TCAATTTTACTGCACTTTTATTCGACGATCCTCCGATTAGCTGCTCCAGTCAACACT 1020
QY 1021 AGAAGAGATGAAACGGCACAATTCGGCTGAAGCTGTCATCGGTTACTTCAGATTTA 1080
DB 1021 AGAAGAGATGAAACGGCACAATTCGGCTGAAGCTGTCATCGGTTACTTCAGATTTA 1080
QY 1081 GAAGGGGATTTCCGATGTTGCTGTTTGGCATTTTCCAAACAGCACAATAAAGGGTTATG 1140
DB 1081 GAAGGGGATTTCCGATGTTGCTGTTTGGCATTTTCCAAACAGCACAATAAAGGGTTATG 1140
QY 1141 TTTTAAATACTACTATTTCGACGATTTGCTGCTAAAGAGGGGTATCTCTCCAGAAA 1200
DB 1141 TTTTAAATACTACTATTTCGACGATTTGCTGCTAAAGAGGGGTATCTCTCCAGAAA 1200
QY 1201 AGAGAGGCTGAAAGCCAGGCCAGGCCAGGCCAGGCCAGGAAAGAAACATGATCATG 1260
DB 1201 AGAGAGGCTGAAAGCTGAAAGCTGAAAGCTGAAAGCTGAAAGCTGAAAGCTGAAAGCT 1260
QY 1261 GCGCCCGGAGAGGATGAAATTTGGCTTCCCGGCTGTCACGCCAGCAGTGACGGAG 1320
DB 1261 GCGCCCGGAGAGGATGAAATTTGGCTTCCCGGCTGTCACGCCAGCAGTGACGGAG 1320
QY 1321 AGAGGTTGCTGTTTGTATGACAGTGTCCGGGATTTCCCGTGTCTTCCACCCCATGGCC 1380
DB 1321 AGAGGTTGCTGTTTGTATGACAGTGTCCGGGATTTCCCGTGTCTTCCACCCCATGGCC 1380
QY 1381 ATCGAAGACACTCAAGAGAGAGATGTCCTTTCTAACTAGTGGCGTAGAATTTCCCTAGGG 1440
DB 1381 ATCGAAGACACTCAAGAGAGAGATGTCCTTTCTAACTAGTGGCGTAGAATTTCCCTAGGG 1440
QY 1441 CGGCCCGGATTAATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTGGGCACTTAGC 1500
DB 1441 CGGCCCGGATTAATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTGGGCACTTAGC 1500
QY 1501 AGAAGACCCGCTCTGTAGATTTCTAATCAAGAGGATGTGAGATGCAATTTGCTGAGAG 1560
DB 1501 AGAAGACCCGCTCTGTAGATTTCTAATCAAGAGGATGTGAGATGCAATTTGCTGAGAG 1560
QY 1561 ATCGAGCTTCAATTTTGTATCTTTTATTTTGTAACTATATATATATATATATATATAT 1620
DB 1561 ATCGAGCTTCAATTTTGTATCTTTTATTTTGTAACTATATATATATATATATATATAT 1620
QY 1621 GTCATTTGTTTCTTCTGCTAGAGCTTCTCTGATCAGCTTCTCTGATCAGCTTCTGATGAA 1680
DB 1621 GTCATTTGTTTCTTCTGCTAGAGCTTCTCTGATCAGCTTCTCTGATCAGCTTCTGATGAA 1680
QY 1681 TATCTTGTGTAGGGGTTTGGGAAATCAATTCGAGTTGATGTTTCTTGTGTTTCTTCC 1740
DB 1681 TATCTTGTGTAGGGGTTTGGGAAATCAATTCGAGTTGATGTTTCTTGTGTTTCTTCC 1740
QY 1741 ACTCCTTTCAGAGTACAGAGATTAAGTGAGAGTTCCGTTTGTGCAAGCTTATCGATAA 1800
DB 1741 ACTCCTTTCAGAGTACAGAGATTAAGTGAGAGTTCCGTTTGTGCAAGCTTATCGATAA 1800
QY 1801 GCTTTAATGCGGTAGTTTATCAGTTAAATTTGCTAAACGAGTCAAGCACCCTGTATGAA 1860
DB 1801 GCTTTAATGCGGTAGTTTATCAGTTAAATTTGCTAAACGAGTCAAGCACCCTGTATGAA 1860
QY 1861 ATCTAAATGCGGCTCATGCTCATCTCCGCAACCGTCAACCTGGATGCTGTAGGCAATG 1920
DB 1861 ATCTAAATGCGGCTCATGCTCATCTCCGCAACCGTCAACCTGGATGCTGTAGGCAATG 1920
QY 1921 CTTGGTTATGCGGTAATGCGGCTCTTGGGGATATCGCCATTCGACAGCATCGC 1980
DB 1921 CTTGGTTATGCGGTAATGCGGCTCTTGGGGATATCGCCATTCGACAGCATCGC 1980
QY 1981 CAGTCACTATGGCGTCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCGGT 2040
DB 1981 CAGTCACTATGGCGTCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCGGT 2040
QY 1981 CAGTCACTATGGCGTCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCGGT 2040
DB 1981 CAGTCACTATGGCGTCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCGGT 2040
```


Db 3995 GTTCGTTATTCAGTAGTCATTTGCCGAGCATCAATATATGGGATTTATACAGAGCAACAG 4054
Qy 4261 TGGAAAGTCAATATACCAACTTTTGGGTCTCAGAAAAAGCATATAACAGTTCTACTACCGC 4320
Db 4055 TGGAAAGTCAATATACCAACTTTTGGGTCTCAGAAAAAGCATATAACAGTTCTACTACCGC 4114
Qy 4321 CATTAGTGAAACTTTTCAAAATCGCCAGTGGAGAGAAAAGGCACACGCGATATAGCAT 4380
Db 4115 CATTAGTGAAACTTTTCAAAATCGCCAGTGGAGAGAAAAGGCACACGCGATATAGCAT 4174
Qy 4381 TAGCGGCAAGATGCAACTTTTCAAAACCGAGGTCTTATAGATAACCCCTAGCGGCTGGGA 4440
Db 4175 TAGCGGCAAGATGCAACTTTTCAAAACCGAGGTCTTATAGATAACCCCTAGCGGCTGGGA 4234
Qy 4441 TCATCCCTTGGACAACTTTTCTGCCAAATCTAGTCCCAAAATCAATCAATCTTCTAGTATCCAT 4500
Db 4235 TCATCCCTTGGACAACTTTTCTGCCAAATCTAGTCCCAAAATCAATCAATCTTCTAGTATCCAT 4294
Qy 4501 TATTGTCAAACTTGAGCAAGTTGTGATCAGCTCTCTCAAAATTTGGTCTCTGTAAACGGATG 4560
Db 4295 TATTGTCAAACTTGAGCAAGTTGTGATCAGCTCTCTCAAAATTTGGTCTCTGTAAACGGATG 4354
Qy 4561 ACTCAACTTGCACATTAATTAAGCTCAGTTCGATTTGAGTGAACCTTGTATCAGGTTGTGCA 4620
Db 4355 ACTCAACTTGCACATTAATTAAGCTCAGTTCGATTTGAGTGAACCTTGTATCAGGTTGTGCA 4414
Qy 4621 GCTGTCAGCAGCATAGGGAACACGGCTTTTCTCAAAACTCAAGGAATTAATCAAACT 4680
Db 4415 GCTGTCAGCAGCATAGGGAACACGGCTTTTCTCAAAACTCAAGGAATTAATCAAACT 4474
Qy 4681 CTGCAACTTTGCGATGACAGGTAGCAAGGAAATGTCAATCTTGAAGTGGACAGTGAG 4740
Db 4475 CTGCAACTTTGCGATGACAGGTAGCAAGGAAATGTCAATCTTGAAGTGGACAGTGAG 4534
Qy 4741 TGTAGTCTTGAGAAATCTGAAGCGGTATTTTATATATCAGTGAATCAGTATCAGGAG 4800
Db 4535 TGTAGTCTTGAGAAATCTGAAGCGGTATTTTATATATCAGTGAATCAGTATCAGGAG 4594
Qy 4801 TCCTCTACGCGGAGCGATCTGTGGCGGA- 4828
Db 4595 TCCTCTACGCGGAGCGATCTGTGGCGGACCTGTGAGGGGGGGGGGGGGGGGGTCTGC 4654
Qy 4829 - 4828
Db 4655 CTCGTGAAGAAGTGTGTGCTGACTCATACAGGCGCTGAATCGCCCCATCATCCAGCAGA 4714
Qy 4829 - 4828
Db 4715 AAGTCAGGAGCCACGGTTGATGAGAGCTTTGTGTAGGTGGACAGTTGATTTGA 4774
Qy 4829 - 4828
Db 4775 ACTTTTGTCTTGGCCACGGAAACGGTCTGCGTTGTGCGGAAGATGGGTGATCTGATCCTTCA 4834
Qy 4829 - 4828
Db 4835 ACTCAGCAAAAGTTCGATTTATTAACAAAGCCGCGCTCCCGTCMAGTCAAGTCAAGTATGCT 4894
Qy 4829 - 4828
Db 4895 CTGCCAGTGTACACCAATTAACCAATTCGATTGAGAAAAAATCATCGAGCATCAAAATG 4954
Qy 4829 - 4828
Db 4955 AAAGTCAATTTATTATATATCAGGATTAATCAATACCATATTTTGAAGAAAGCCGTTTCTG 5014
Qy 4829 - 4828
Db 5015 TAAATGAAGAGAAACTCACCGGAGCTTCCATAGGATGGCAAGATCCTGGTATCGGTC 5074
Qy 4829 - 4828

Db 5075 TCGGATTCGACTCGTCCAAACATCAATCAACACCTTATAATTTTCCCTCTGCTCAAAATAAG 5134
Qy 4829 - 4828
Db 5135 GTTATCAAGTGAATAATCAACATGAGTGAAGTGAATCCGGTGAGATGGCAAGCTT 5194
Qy 4829 - 4828
Db 5195 ATGCATTTCTTCCAGACTTTGTTCAACAGGCGACCATTTACGCTCGTCAATCAAAATCACT 5254
Qy 4829 - 4828
Db 5255 CGCATCAACCAACCGGTTATTCATTCGTGATTGCGCTGAGCGAGACGAAATACGCGATC 5314
Qy 4829 - 4828
Db 5315 GCTGTTAAAGGACAAATTAACAAACAGGAATCGAATGCAACCGCGAGCAACACATGCCAG 5374
Qy 4829 - 4828
Db 5375 CGCATCAACAAATTTTCACTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTT 5434
Qy 4829 - 4828
Db 5435 CCGGGGATCGCAGTGGTGAATACCATGCATCATCAGGATACGGATAAAATGCTTGAT 5494
Qy 4829 - 4828
Db 5495 GGTGCGAAGAGGCATAAATCCGTGAGCAGTTTAGTCTGACCAATCTCATCTGAACATC 5554
Qy 4829 - 4828
Db 5555 ATTGGAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCAT 5614
Qy 4829 - 4828
Db 5615 CAATCGATAGATTGTCGACCTGATTGCGGACATTTATCGGAGCCCAATTTATACCCATA 5674
Qy 4829 - 4828
Db 5675 TAAATCAGCATCCATGTTGGAATTTAATCGCGCTCGAGCAAGAGTTTCCCGTTGAAT 5734
Qy 4829 - 4828
Db 5735 ATGGTCAATAACCCCTTGTATTACTGTTTATGTAAAGCAGACAGTTTATTGTTTCATGA 5794
Qy 4829 - 4828
Db 5795 TGATATATTTTATCTTTGCAATGTAAACATCAGAGATTTTGAGACACACGTTGCTTTC 5854
Qy 4829 - 4828
Db 5855 CCCCCCCCCCTCGAGTGGGATACCGCGCCACAGGTGCGGTTGCTGGCGCTTATAT 5914
Qy 4880 CGCGGACATCACCGATGGGAAAGATCGGCTCGCCACTTCGCGCTCATAGCGCTTGTCTT 4939
Db 5915 CGCGGACATCACCGATGGGAAAGATCGGCTCGCCACTTCGCGCTCATAGCGCTTGTCTT 5974
Qy 4940 CGCGGTTGGTATGTTGGCAGGCGGCTGCGCGGCGGAGCTGTTGGGCGCCATCTCCTTGA 4999
Db 5975 CGCGGTTGGTATGTTGGCAGGCGGCTGCGCGGCGGAGCTGTTGGGCGCCATCTCCTTGA 6034
Qy 5000 TGCAACATTTCTTGGCGGCGGCTGCTCAACGCGCTCAACCTACTACTGCGGCTGCTTCT 5059
Db 6035 TGCAACATTTCTTGGCGGCGGCTGCTCAACGCGCTCAACCTACTACTGCGGCTGCTTCT 6094
Qy 5060 AATGCAAGGATCGCATTAAGGAGAGCGTGAAGTATCTATGTTGGAAGTATGGGAATGCT 5119
Db 6095 AATGCAAGGATCGCATTAAGGAGAGCGTGAAGTATCTATGTTGGAAGTATGGGAATGCT 6154
Qy 5120 GATACCGGATTTCTCAGTGTCTTGAAGTCTCTATCAGATATGCGCAATTAAGCAAC 5179
Db 6155 GATACCGGATTTCTCAGTGTCTTGAAGTCTCTATCAGATATGCGCAATTAAGCAAC 6214


```

8375 TAAGTAGTCCCGCAGTTAATAGTTGGCCAAAGTTGTGCGCATTGCTGCGAGGCATCGTGG 8434
8435 TGTACCGCTCGTCGTTTGGTATGGCTTCATTACAGCTCCGGTTCCTCCAAACGATCAAGGCGAG 7459
8435 TGTACCGCTCGTCGTTTGGTATGGCTTCATTACAGCTCCGGTTCCTCCAAACGATCAAGGCGAG 8494
7460 TTACATGATCCCGCATGTTGTGCAAAAAGCGGTAGTCTCTCGGTCTCCGATCGTTG 7519
8495 TTACATGATCCCGCATGTTGTGCAAAAAGCGGTAGTCTCTCGGTCTCCGATCGTTG 8554
7520 TCAGAAAGTAAGTGGCCGACGTTTATCAGTCATGTTTATGGCAGCACTGCATATTTCTC 7579
8555 TCAGAAAGTAAGTGGCCGACGTTTATCAGTCATGTTTATGGCAGCACTGCATATTTCTC 8614
7580 TTACTGTCATGCCATCCCTAGATGCTTTTCTGTGACGCTGGTGTAGTACTCAACCAAGTCAT 7639
8615 TTACTGTCATGCCATCCCTAGATGCTTTTCTGTGACGCTGGTGTAGTACTCAACCAAGTCAT 8674
7640 TCTGAGATAGTGTATGGCGGACCGAGTGTCTTTCGCCGCGCTCAACCGGATATA 7699
8675 TCTGAGATAGTGTATGGCGGACCGAGTGTCTTTCGCCGCGCTCAACCGGATATA 8734
7700 CCGGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAGCGTTCTTTCGGGGCGAA 7759
8735 CCGGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAGCGTTCTTTCGGGGCGAA 8794
7760 AACTCTCAAGGATCTTACCGCTGTGTGAGATCCAGTTCGATGCACTGTCACCCCA 7819
8795 AACTCTCAAGGATCTTACCGCTGTGTGAGATCCAGTTCGATGCACTGTCACCCCA 8854
7820 ACTGATCTTCAGCATCTTTTACTTTCACAGCGTTTCTGGGTGAGCAAAAACAGGAAGGC 7879
8855 ACTGATCTTCAGCATCTTTTACTTTCACAGCGTTTCTGGGTGAGCAAAAACAGGAAGGC 8914
7880 AAATGTCGCAAAAAGGGAATAAGGCGACACGGAATGTTGTAATCTATCTCTTCC 7939
8915 AAATGTCGCAAAAAGGGAATAAGGCGACACGGAATGTTGTAATCTATCTCTTCC 8974
7940 TTTTTCATATTTATGAGCATTTATCAGGTTTATGTTCTCATGAGCGGATACATATTTG 7999
8975 TTTTTCATATTTATGAGCATTTATCAGGTTTATGTTCTCATGAGCGGATACATATTTG 9034
8000 AATGTTATTTAGAAAATAAACAATAGGGTTTCGCGCACATTTCCCGAAAAGTGCCAC 8059
9035 AATGTTATTTAGAAAATAAACAATAGGGTTTCGCGCACATTTCCCGAAAAGTGCCAC 9094
8060 CTGAGCTCTAAGAACCATTTATTCATGACATTAACCTATAAATAAGGCGTATCAGA 8119
9095 CTGAGCTCTAAGAACCATTTATTCATGACATTAACCTATAAATAAGGCGTATCAGA 9154
8120 GGCCCTTTCGTTCTCAAGAAATTAATTTCTCATGTTTGACAGCTTATCATCGATAAGCTGAC 8179
9155 GGCCCTTTCGTTCTCAAGAAATTAATTTCTCATGTTTGACAGCTTATCATCGATAAGCTGAC 9214
8180 TCATGTTGGTATTTGTAATAGACGAGATCGGGAACCTGMAAATAACAGTTATTTATT 8239
9215 TCATGTTGGTATTTGTAATAGACGAGATCGGGAACCTGMAAATAACAGTTATTTATT 9274
8240 CG 8241
9275 CG 9276

```

RESULT 9

ABZ22459

ID ABZ22459 standard; DNA; 9276 BP.

XX ABZ22459;

AC ABZ22459;

XX 24-MAR-2003 (first entry)

XX Pichia pastoris vector pPIC9K comprising MTSP10 DNA SEQ ID NO:13.

DE

XX Human; membrane-type serine protease; enzyme; MTSP10; cytostatic;
 KW type-II membrane-type serine protease; neoplastic disease; tumour; gene;
 OS da.

XX Pichia pastoris.

OS Homo sapiens.

OS Synthetic.

XX WO200292841-A2.

XX 21-NOV-2002.

XX 14-MAY-2002; 2002WO-US015332.

XX 14-MAY-2001; 2001US-0291001P.

XX (CORV-) CORVAS INT INC.

XX Madison EL, Yeh J;

XX WPI; 2003-129309/12.

XX New polypeptides comprising the protease domain of a type-II membrane-

PT type serine protease (MTSP10), or its muteins, useful for diagnosing

PT neoplasms or malignancies, or for screening for MTSP10 inhibitors for

PT treating such diseases.

XX Example 2; Page 189-191; 198pp; English.

XX The present invention describes a polypeptide comprising a purified

CC single or two chain polypeptide, which comprises the protease domain of a

CC type-II membrane-type serine protease (MTSP10) or its catalytically

CC active portion, or a mutain of it, where up to 50 % of the amino acids

CC are replaced with another amino acid, and the resulting polypeptide is a

CC single chain or two chain polypeptide that has a catalytic activity of at

CC least 1-10 % of the unmutated polypeptide. MTSP10 has cytostatic

CC activity. The polypeptide containing the protease domain of the MTSP10 is

CC useful for detecting a neoplastic disease, and for diagnosing the

CC presence of a pre-malignant lesion, a malignancy, or other pathologic

CC condition in a subject, or monitoring tumour (e.g. breast, cervix,

CC prostate, lung, ovary or colon tumour) progression and/or therapeutic

CC effectiveness. An inhibitor of the polypeptide containing the protease

CC domain of MTSP10 is useful for treating or preventing neoplastic disease

CC in a mammal. An inhibitor of the activation cleavage of the zymogen form

CC of the MTSP10 polypeptide is useful for inhibiting tumour initiation,

CC growth or progression, or treating (pre-)malignant conditions of the e.g.

CC breast, cervix, prostate, lung, ovary or colon. The present sequence

CC represents a Pichia pastoris vector designated pPIC9K, comprising MTSP10,

CC which is used in an example from the present invention

XX Sequence 9276 BP; 2493 A; 2197 C; 2078 G; 2508 T; 0 U; 0 Other;

XX Query Match 79.6%; Score 6561.6; DB 7; Length 9276;

XX Best Local Similarity 84.7%; Pred. No. 0;

XX Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;

Qy 1 AGATCTAACATCCAAAGACGAAAGGTTGATGAAACCTTTTGGCCATCCGACATCCACAG 60

Db 1 AGATCTAACATCCAAAGACGAAAGGTTGATGAAACCTTTTGGCCATCCGACATCCACAG 60

Qy 61 GTCCATTTCTCACATAGTGCACAAACGCAACAGAGGGGATACACTAGCAGACAGCCGT 120

Db 61 GTCCATTTCTCACATAGTGCACAAACGCAACAGAGGGGATACACTAGCAGACAGCCGT 120

Qy 121 TGCAAAACGAGGAGCTCCACT 180

Db 121 TGCAAAACGAGGAGCTCCACT 180

Qy 181 AGCCAGTTATTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 240

Db 181 AGCCAGTTATTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 240

QY 241 ACACATGACCTTTATTAGCCTGTCTATCCCTGGCCCCCTCGCGAGGTTTCATGTTTGTATA 300
DB 241 ACACATGACCTTTATTAGCCTGTCTATCCCTGGCCCCCTCGCGAGGTTTCATGTTTGTATA 300
QY 301 TTTCCGAATGCAACAAGCTCCGCTATTACACCGCAACATCACTCCAGATGAGGCTTTCTG 360
DB 301 TTTCCGAATGCAACAAGCTCCGCTATTACACCGCAACATCACTCCAGATGAGGCTTTCTG 360
QY 361 AGTGTGGGCTCAATATGTTTCATGTTCCCAATGGCCCAAACTGACAGTTTAAAGCT 420
DB 361 AGTGTGGGCTCAATATGTTTCATGTTCCCAATGGCCCAAACTGACAGTTTAAAGCT 420
QY 421 GTCCTTGGAACTTAATATGACAAAAGCGTGTCTCATCCAAAGATGAACCTAAGTTTGGTTCC 480
DB 421 GTCCTTGGAACTTAATATGACAAAAGCGTGTCTCATCCAAAGATGAACCTAAGTTTGGTTCC 480
QY 481 TTGAATGCTAAGCGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCCGCCATACCGTTTGT 540
DB 481 TTGAATGCTAAGCGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCCGCCATACCGTTTGT 540
QY 541 CTGTTTGGTATGATTTGATGCAAGAACTGCTCAAAAATAAATCTCATTAATGCTTAGCGCAGTCT 600
DB 541 CTGTTTGGTATGATTTGATGCAAGAACTGCTCAAAAATAAATCTCATTAATGCTTAGCGCAGTCT 600
QY 601 CTCTATCGCTTCTGAACCCCGGTGCACTGTGCGGAAACGCAATGGGAAACACCCGCT 660
DB 601 CTCTATCGCTTCTGAACCCCGGTGCACTGTGCGGAAACGCAATGGGAAACACCCGCT 660
QY 661 TTTTGGATGATTGATGCTCCCAATGCTATGCTTCCAAAGATTCGTGGGGAATACT 720
DB 661 TTTTGGATGATTGATGCTCCCAATGCTATGCTTCCAAAGATTCGTGGGGAATACT 720
QY 721 GCTGATAGCTTAAGCTTCATGATCAAAATTTAACTGTTCTAAACCCCTACTTGACAGCAAT 780
DB 721 GCTGATAGCTTAAGCTTCATGATCAAAATTTAACTGTTCTAAACCCCTACTTGACAGCAAT 780
QY 781 ATATAAACAGAAAGAGTGGCTGCTTAAACCTTTTATTAATCATCATTAATAGCTT 840
DB 781 ATATAAACAGAAAGAGTGGCTGCTTAAACCTTTTATTAATCATCATTAATAGCTT 840
QY 841 ACTTTTCAATTTGGACCTGGTTCCAAATGCAAGCTTTGATTTTAAACGACTTTTAAACGA 900
DB 841 ACTTTTCAATTTGGACCTGGTTCCAAATGCAAGCTTTGATTTTAAACGACTTTTAAACGA 900
QY 901 CAACCTTGAGAGATCAAAAAACAATTAATTCGAAGGATCCAAACGATGAGATTTCCCT 960
DB 901 CAACCTTGAGAGATCAAAAAACAATTAATTCGAAGGATCCAAACGATGAGATTTCCCT 960
QY 961 TCAATTTTACTGCGAGTTTATTCGAGCATCTCCGCAATTAGCTGCTCCAGTCAACACT 1020
DB 961 TCAATTTTACTGCGAGTTTATTCGAGCATCTCCGCAATTAGCTGCTCCAGTCAACACT 1020
QY 1021 ACAACAGAAAGATGAAACGGCAAAATTCGCGCTGAAGCTGTCATCGGTTACTCAGATTTA 1080
DB 1021 ACAACAGAAAGATGAAACGGCAAAATTCGCGCTGAAGCTGTCATCGGTTACTCAGATTTA 1080
QY 1081 GAAGGGATTTGATGTTGCTGTTTGGCATTTTCCACAGCACAATAAGCGGTTATTG 1140
DB 1081 GAAGGGATTTGATGTTGCTGTTTGGCATTTTCCACAGCACAATAAGCGGTTATTG 1140
QY 1141 TTTTATAATATCTATTTGCCAGATTTGCTGCTAAAGAAAGGGGTATCTCTCGAGAAA 1200
DB 1141 TTTTATAATATCTATTTGCCAGATTTGCTGCTAAAGAAAGGGGTATCTCTCGAGAAA 1200
QY 1201 AGAGAGGCTGAAGCCCGGCGCCAGGCCCAAGGCCCAAGGAAACAATGATCATG 1260
DB 1201 AGAGAGGCTGAAGCTTA----- 1217
QY 1261 GCCCCCCGGGAGAGGATAAATTTGGCTTCCCGGTGTACCCGCTGACCCGACGAGTGCACGGAG 1320
DB 1218 ----- 1217
QY 1321 AGAGGTTGCTGTTTGTATGACAGTGTCCGGGATTCGGTGGTCTCCACCCCATGGCC 1380

DB 1218 ----- 1217
QY 1381 ATCGAGAACACTCAAGAAAGAAATGTCCTTCTAACTAGTGGCGTAGAATTCCTTAGGG 1440
DB 1218 -----CGTAGAATTCCTTAGGG 1234
QY 1441 CGCGCCGGAATTTAATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTGGGCACCTTACG 1500
DB 1235 CGCGCCGGAATTTAATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTGGGCACCTTACG 1294
QY 1501 AGAAGACCGGCTTCTAGATCTTAATCAAGAGAGTGCAGAAATGCCATTTGCGCTGAGAG 1560
DB 1295 AGAAGACCGGCTTCTAGATCTTAATCAAGAGAGTGCAGAAATGCCATTTGCGCTGAGAG 1354
QY 1561 ATGCAAGGCTTCATTTTTCATCTTTTATTTGTAACCTATATAGTATAGGATTTTCTT 1620
DB 1355 ATGCAAGGCTTCATTTTTCATCTTTTATTTGTAACCTATATAGTATAGGATTTTCTT 1414
QY 1621 GTCATTTTGTCTTCTCGTAGAGCTTCTCTCATCAGCCTATCTCGCAGCTGATGAA 1680
DB 1415 GTCATTTTGTCTTCTCGTAGAGCTTCTCTCATCAGCCTATCTCGCAGCTGATGAA 1474
QY 1681 TATCTTGTGTAGGGTTTGGGAAATCATTCGAGTTTGTGTTTCTTCTGATTTTCCC 1740
DB 1475 TATCTTGTGTAGGGTTTGGGAAATCATTCGAGTTTGTGTTTCTTCTGATTTCCC 1534
QY 1741 ACTCTCTCTTCAAGATACAGAAATTAAGTGAGAAAGTTCTGTTTGTGCAAGCTTATCGATA 1800
DB 1535 ACTCTCTCTTCAAGATACAGAAATTAAGTGAGAAAGTTCTGTTTGTGCAAGCTTATCGATA 1594
QY 1801 GCTTTAATTCGGTAGTTTATCACAGTTAAATGCTAACGCACTCAGGCACCGTGTATGAA 1860
DB 1595 GCTTTAATTCGGTAGTTTATCACAGTTAAATGCTAACGCACTCAGGCACCGTGTATGAA 1654
QY 1861 ATCTAACCAATGGCTCATGCTCATCTCGCACCGTCAACCTGGATGCTGTAGGCATAGG 1920
DB 1655 ATCTAACCAATGGCTCATGCTCATCTCGCACCGTCAACCTGGATGCTGTAGGCATAGG 1714
QY 1921 CTTGTTTATGCGGTACTGCGCGGCTCTTGGGGATATGCTCCATTCGACAGATGCG 1980
DB 1715 CTTGTTTATGCGGTACTGCGCGGCTCTTGGGGATATGCTCCATTCGACAGATGCG 1774
QY 1981 CAGTCACTATGCGGTGCTAGCGCTATATGCTGTGATGCAATTTCTATCGCACCCGT 2040
DB 1775 CAGTCACTATGCGGTGCTAGCGCTATATGCTGTGATGCAATTTCTATCGCACCCGT 1834
QY 2041 TCTCGAGACACTGTCCGACCGCTTTGGCGCGCCCGCTCTGCTGCTCTGCTACTTGG 2100
DB 1835 TCTCGAGACACTGTCCGACCGCTTTGGCGCGCCCGCTCTGCTGCTCTGCTACTTGG 1894
QY 2101 AGCCATATCGACTAGCGATCATGGCGACACACCGCTGCTGATCGATCTGATCTA 2160
DB 1895 AGCCATATCGACTAGCGATCATGGCGACACACCGCTGCTGATCGATCTGATCGATCTA 1954
QY 2161 AATGTAAGTTAAATCTCTAAATTAATTAATTAAGTCCAGTTTCTCCATACGAACTTAA 2220
DB 1955 AATGTAAGTTAAATCTCTAAATTAATTAATTAAGTCCAGTTTCTCCATACGAACTTAA 2014
QY 2221 CAGCATTCGGGTGAGCATCTAGACCTTCAAACGAGCGCAGATCCATCTACTGTTGGCCAA 2280
DB 2015 CAGCATTCGGGTGAGCATCTAGACCTTCAAACGAGCGCAGATCCATCTACTGTTGGCCAA 2074
QY 2281 TATGTTTCACTCCCTCAGGATTTAGCTCTTGTGAAGTGAATGAACTTCTGGAAGGTTGAG 2340
DB 2075 TATGTTTCACTCCCTCAGGATTTAGCTCTTGTGAAGTGAATGAACTTCTGGAAGGTTGAG 2134
QY 2341 TGTAACTCCGCTGATTTGACGGGCATATCCGTACGTTGGCAAGATGTGTTGTGATCCGG 2400
DB 2135 TGTAACTCCGCTGATTTGACGGGCATATCCGTACGTTGGCAAGATGTGTTGTGATCCGG 2194
QY 2401 AGAGGTTAATCTCCCAACTCTCTCGAGAGTAGGCACCAACAAACACAGATCCAGCGTGT 2460

Db 2195 AGGAGTAATCTCCAAACTCTCTGGAGAGTAGGCACCAACAAACACAGATCCAGCGTGT 2254
Qy 2461 GTACTGATCAACATAGAGAAGCATTCTCGATTTTCAGAGATCAAGTGTTCAGGAGCGT 2520
Db 2255 GTACTGATCAACATAGAGAAGCATTCTCGATTTTCAGAGATCAAGTGTTCAGGAGCGT 2314
Qy 2521 ACTGATTGGACATTTCCAAAGCCTGCTCGTAGGTGCAACCGATAGGAGTGTAGAGTGTG 2580
Db 2315 ACTGATTGGACATTTCCAAAGCCTGCTCGTAGGTGCAACCGATAGGAGTGTAGAGTGTG 2374
Qy 2581 CAATACACTTCGGTACAAATTTCAACCCCTGGCAACTGCACAGCTTGGTGTGAAACAGCAT 2640
Db 2375 CAATACACTTCGGTACAAATTTCAACCCCTGGCAACTGCACAGCTTGGTGTGAAACAGCAT 2434
Qy 2641 CTTCAATTTCTGGCAAGCTTCTGTGTCATATCGACAGCCCAACAGAAATCACCTGGGAAT 2700
Db 2435 CTTCAATTTCTGGCAAGCTTCTGTGTCATATCGACAGCCCAACAGAAATCACCTGGGAAT 2494
Qy 2701 CAATACCAATGTTAGCTTGGACAGAGAGTCTGAGGCAACGAATCTGGATCAGCGTATT 2760
Db 2495 CAATACCAATGTTAGCTTGGACAGAGAGTCTGAGGCAACGAATCTGGATCAGCGTATT 2554
Qy 2761 TATCAGCAATTAACATAGAACTTTCAGAAAGCCAGAGGATGTCATATCTACACAGGCGTG 2820
Db 2555 TATCAGCAATTAACATAGAACTTTCAGAAAGCCAGAGGATGTCATATCTACACAGGCGTG 2614
Qy 2821 ATGTGTCATTTGAAACCAATCATTTGTCAGAGTAACGAACTGGTTCCTGGACCAAAATA 2880
Db 2615 ATGTGTCATTTGAAACCAATCATTTGTCAGAGTAACGAACTGGTTCCTGGACCAAAATA 2674
Qy 2881 TTTTGTACACTTAGGACAGTTCCTGTCGTAGCCATAGCAGCTACTCCCTGGGCGC 2940
Db 2675 TTTTGTACACTTAGGACAGTTCCTGTCGTAGCCATAGCAGCTACTCCCTGGGCGC 2734
Qy 2941 CTCCTGCTAGCAGATACACTTAGCACCACCTGTTGGGCAACGTAGATGACTTCTGGGG 3000
Db 2735 CTCCTGCTAGCAGATACACTTAGCACCACCTGTTGGGCAACGTAGATGACTTCTGGGG 2794
Qy 3001 TAAGGGTACCATCTTTAGTGGAGATGCAAAACAAATTTCTTTGCAACCAAGCAACTT 3060
Db 2795 TAAGGGTACCATCTTTAGTGGAGATGCAAAACAAATTTCTTTGCAACCAAGCAACTT 2854
Qy 3061 TGGCAGGAACACCCAGCATCAGGAAAGTGGAAAGGAGAAATTCGGGTTCCACAGGAATAT 3120
Db 2855 TGGCAGGAACACCCAGCATCAGGAAAGTGGAAAGGAGAAATTCGGGTTCCACAGGAATAT 2914
Qy 3121 AGAGCCCAACTTTCTCAATAGCTTTGCAAAACGAGAGACACTACACAGGCAAGTCT 3180
Db 2915 AGAGCCCAACTTTCTCAATAGCTTTGCAAAACGAGAGACACTACACAGGCAAGTCT 2974
Qy 3181 CAACTTGCAACGCTCCGTTAGTTCAGCTTCATGGAATTCCTGAGCTTATCTATAGAGA 3240
Db 2975 CAACTTGCAACGCTCCGTTAGTTCAGCTTCATGGAATTCCTGAGCTTATCTATAGAGA 3034
Qy 3241 GATCAATGGCTCTCTTAACGTTATCTGGCAATTCGCAATTCCTCTGGGAAAGGAGCTT 3300
Db 3035 GATCAATGGCTCTCTTAACGTTATCTGGCAATTCGCAATTCCTCTGGGAAAGGAGCTT 3094
Qy 3301 CTAACACAGGTGCTTCAAGAGGATCCATCAAACTGGCAGTTCATTAAGGGCTT 3360
Db 3095 CTAACACAGGTGCTTCAAGAGGATCCATCAAACTGGCAGTTCATTAAGGGCTT 3154
Qy 3361 TGTCAACATTTGACGAACATTTGCAAAATTTGGTTCGACTAATTCCTGCAATTCCTGCG 3420
Db 3155 TGTCAACATTTGACGAACATTTGCAAAATTTGGTTCGACTAATTCCTGCAATTCCTGCG 3214
Qy 3421 TTTTCTGGATAGGACGAAGGCAATTTCAATTTCTTGTGAGGAGGCTTAGAAAGCT 3480
Db 3215 TTTTCTGGATAGGACGAAGGCAATTTCAATTTCTTGTGAGGAGGCTTAGAAAGCT 3274
Qy 3481 CAAATTTGCAATTCATACGACCTTTCAGAAAGGAGCTTCTTTAGTTTGGATCTTCTT 3540
Db 3275 CAAATTTGCAATTCATACGACCTTTCAGAAAGGAGCTTCTTTAGTTTGGATCTTCTT 3334

Qy 3541 TAGGTTGTTCTCTGGTGTATCTCGCTTGGCATCTCTTTCTCTCTAGTGACCTTTAGGG 3600
Db 3335 TAGGTTGTTCTCTGGTGTATCTCGCTTGGCATCTCTTTCTCTCTAGTGACCTTTAGGG 3394
Qy 3601 ACTTCATATCCAGGTTTCTCTCCACTCGTCAAGTGCACACCGTACTTGGCACAATCTAA 3660
Db 3395 ACTTCATATCCAGGTTTCTCTCCACTCGTCAAGTGCACACCGTACTTGGCACAATCTAA 3454
Qy 3661 CTAATGCAAAATAAATAAGTCCAGCACAATTCOCAGGCTATATCTCTTTGGATTTAGCTT 3720
Db 3455 CTAATGCAAAATAAATAAGTCCAGCACAATTCOCAGGCTATATCTCTTTGGATTTAGCTT 3514
Qy 3721 CTGCAAGTTCATCAGCTTCTCTCCCTAAATTTTAGCGTTCAACAAAATTCGTCGTAATA 3780
Db 3515 CTGCAAGTTCATCAGCTTCTCTCCCTAAATTTTAGCGTTCAACAAAATTCGTCGTAATA 3574
Qy 3781 ACCGTTTGTATAAGAACCTTCTGGAGCATTCGCTTACGATCCCAAGGTGGCTTCCA 3840
Db 3575 ACCGTTTGTATAAGAACCTTCTGGAGCATTCGCTTACGATCCCAAGGTGGCTTCCA 3634
Qy 3841 TGGCTCTAAGACCCCTTTGATTTGGCCAAAACAGGAGTGGCTTCCAAGTGCAGAAAACCA 3900
Db 3635 TGGCTCTAAGACCCCTTTGATTTGGCCAAAACAGGAGTGGCTTCCAAGTGCAGAAAACCA 3694
Qy 3901 CACCTGTTTGTCAACCCACAATTTTCAAGCAGTCTCCATCAATTCGAATTCGATACCCA 3960
Db 3695 CACCTGTTTGTCAACCCACAATTTTCAAGCAGTCTCCATCAATTCGAATTCGATACCCA 3754
Qy 3961 GCAACTTTTGTAGTGTCTCAGATGTAGCACCTTTATACCAAAACCGTGCAGACGAGAT 4020
Db 3755 GCAACTTTTGTAGTGTCTCAGATGTAGCACCTTTATACCAAAACCGTGCAGACGAGAT 3814
Qy 4021 GGTAGACTCCAGTTTGTCTTATAGCTTCCGGAATAGACTTTTGGACGAGTACACCA 4080
Db 3815 GGTAGACTCCAGTTTGTCTTATAGCTTCCGGAATAGACTTTTGGACGAGTACACCA 3874
Qy 4081 GGCACCAAGTAAATTAGAAGAGTGCAGCCACCAAGTAGTGAATAGACCATTCGGGGCGGT 4140
Db 3875 GGCACCAAGTAAATTAGAAGAGTGCAGCCACCAAGTAGTGAATAGACCATTCGGGGCGGT 3934
Qy 4141 CAGTAGTCAAGAGCGCAACAAATTTCTAGCAGGGAACCTTTTGGACATCTTTCAGAAA 4200
Db 3935 CAGTAGTCAAGAGCGCGCAACAAATTTCTAGCAGGGAACCTTTTGGACATCTTTCAGAAA 3994
Qy 4201 GTTCGTAATTCAGTAGTCAATTCGCGAGCATCAATATGGGGATTATACCAAGAACACAG 4260
Db 3995 GTTCGTAATTCAGTAGTCAATTCGCGAGCATCAATATGGGGATTATACCAAGAACACAG 4054
Qy 4261 TGGAGTCAATCTACCAACTTTGCGGTCTCAGAAAAAGCATAAAACAGTTCTACTACCGC 4320
Db 4055 TGGAGTCAATCTACCAACTTTGCGGTCTCAGAAAAAGCATAAAACAGTTCTACTACCGC 4114
Qy 4321 CATTAGTGAACCTTTTCAATCCCGCAGTGGAGAAAAGGACAGCGATACTAGCAT 4380
Db 4115 CATTAGTGAACCTTTTCAATCCCGCAGTGGAGAAAAGGACAGCGATACTAGCAT 4174
Qy 4381 TAGCGGCAAGGATGCAACTTTTATCAACAGGCTCTCTATAGATAACCTTAGCGCTGGGA 4440
Db 4175 TAGCGGCAAGGATGCAACTTTTATCAACAGGCTCTCTATAGATAACCTTAGCGCTGGGA 4234
Qy 4441 TCATPCTTTGGACAACTCTTTTTCGCAAAATCTAGGTCGCAAAATCAGTTCTGATACCAT 4500
Db 4235 TCATPCTTTGGACAACTCTTTTTCGCAAAATCTAGGTCGCAAAATCAGTTCTGATACCAT 4294
Qy 4501 TATTGTCAAACTTTGAGCAAGTTCGATCAGCTCTCTCAAAATTCGTTCTCTGTAACGGATG 4560
Db 4295 TATTGTCAAACTTTGAGCAAGTTCGATCAGCTCTCTCAAAATTCGTTCTCTGTAACGGATG 4354
Qy 4561 ACTCAACTTCGACATTAACCTTGAAGCTCAGTTCGATTCGATTCAGTTGAGTTCGCA 4620
Db 4355 ACTCAACTTCGACATTAACCTTGAAGCTCAGTTCGATTCGATTCGATTCAGTTGAGTTCGCA 4414

4621 GCTGTCAGCAGCATAGGGAACACGCGTTTTCCTACCAAACTCAAGGAATTTCAAACT 4680
4415 GCTGTCAGCAGCATAGGGAACACGCGTTTTCCTACCAAACTCAAGGAATTTCAAACT 4474
4681 CTGCAACACTTTCCTATCCAGGTAGCAAGGAATGTCATCTGAACTCGGACAGTGAG 4740
4475 CTGCAACACTTTCCTATCCAGGTAGCAAGGAATGTCATCTGAACTCGGACAGTGAG 4534
4741 TGTAGTCTTGAGAAATCTGAAAGCGGTATTTTATTATCACTGAGTCAGTCATCAGGAGA 4800
4535 TGTAGTCTTGAGAAATCTGAAAGCGGTATTTTATTATCACTGAGTCAGTCATCAGGAGA 4594
4801 TCCTCTAGCCGGAGCGCATCTGTGGCGA----- 4828
4595 TCCTCTAGCCGGAGCGCATCTGTGGCGACCTGCAGGGGGGGGGGGCGCTGAGGTCGC 4654
4829 ----- 4828
4655 CTCGTGAAGAGGTGTTGCTGACTCATACAGGCGCTGAATCGCCCCCATCATCAGCCAGA 4714
4829 ----- 4828
4715 AAGTGAGGAGCCAGGTTGATGAGAGCTTTGTTGTAGGTGACACAGTGTGTGATTTTGA 4774
4829 ----- 4828
4775 ACTTTGCTTTGCCACGGAACGGTCTGCGTTGTCTGGGAAGATGCGTGATCTGATCCTTCA 4834
4829 ----- 4828
4835 ACTCAGCAAAAGTTCGATTTATTCAACAAAGCGCGCTCCCGTCAAGTCAGCGGTAATGCT 4894
4829 ----- 4828
4895 CTGCCAGTGTACAAACCAATTAACCAATCTGATTAGAAAAAATCTACGAGCATCAAATG 4954
4829 ----- 4828
4955 AAACGTCAATTTATCATATCAGGTTATCAATACCATATTTTGTGAAAAGCCGTTTCG 5014
4829 ----- 4828
5015 TAATGAAGAGAAAACTCACCGAGGCGAGTTCCATPAGGATGGCAAGATCCTGGTATCGGTC 5074
4829 ----- 4828
5075 TGGGATTCGGAATCGTCCACATCAATACACCTATTAATTTCCCTCGTCAAAAATAAG 5134
4829 ----- 4828
5135 GTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAAATGGCAAAAGCTT 5194
4829 ----- 4828
5195 ATGATTTCTTCCAGACTTGTTCACAGGCCAGCCATTAAGTTCGCTGATCAAAATCACT 5254
4829 ----- 4828
5255 CGCATCAACCAACCGTTATTTCATTTCGTTGCTGCGCTGAGCGAGAGCAAAATACGCGATC 5314
4829 ----- 4828
5315 GCTGTTAAAGGACAAATTAACAAAGGAATCGAAATGCAACCGGCGCAGGAACACTGCGCAG 5374
4829 ----- 4828
5375 CGCATCAACAAATTTTCACTGAGTACAGGATATTTCTTCAATACCTGGAATGCTGTTTT 5434
4829 ----- 4828
5435 CCCGGGATCGAGTGTGAGTAACCATGCAATCATCAGGAGTACGGATAAAATGCTTTGAT 5494
4829 ----- 4828

5495 GCTCGAAGAGGCAATAAATTCGGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACAATC 5554
4829 ----- 4828
5555 ATTGCAACGCTACCTTTGCCATGTTTCAGAAACAACACTCTGCGCATCGGCTTCCCATATA 5614
4829 ----- 4828
5615 CAATCGATAGATTGTGCAACCTGATTGCCGACATTTATCGGAGCCCAATTTATACCCATA 5674
4829 ----- 4828
5675 TAAATCAGCATCCATGTTGGAAATTTAATCGCGGCTCGAGCAAGACGTTTCCCGTTGAAAT 5734
4829 ----- 4828
5735 ATGGCTCATACACACCCCTTGTATTACTGTATTATGTAAGACACAGTTTTATTGTTTCATGA 5794
4829 ----- 4828
5795 TGATATATTTTATCTTGTGCAATGTAAATCATCAGAGATTTTGAGACACAACGTCGCTTTC 5854
4829 ----- 4828
5855 CCCCCCCCCCTGCAAGTTCGCAATCACCGCGCCACACAGTTCGGTTCGCGGCTATAT 5914
4880 CCGCAGACATCCAGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGGCTTGTTT 4939
5915 CGCCGACATCAACGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGGCTTGTTT 5974
4940 CCGGCTGGGTATGTTGGCAGGCCCCGTCGGCGGGGACTGTTGGCGCCATCTCTCTGCA 4999
5975 CCGGCTGGGTATGTTGGCAGGCCCCGTCGGCGGGGACTGTTGGCGCCATCTCTCTGCA 6034
5000 TCACCAATTCCTTGGCGGGGCTCTCAACGGGCTCAACCTACTACTCTGCGGCTGCTTCT 5059
6035 TGACCAATTCCTTGGCGGGGCTCTCAACGGGCTCAACCTACTACTCTGCGGCTGCTTCT 6094
5060 AATGAGAGTGCATTAAGGAGAGCGTCGAGTATCTATGATTTGAAATGGAAGTATGGAATG 5119
6095 AATGAGAGTGCATTAAGGAGAGCGTCGAGTATCTATGATTTGAAATGGAAGTATGGAATG 6154
5120 GATACCCGCTTCTTCAAGTCTTCAAGTCTCTATCAAGTATGATGATGATGATGATGATGAT 5179
6155 GATACCCGCTTCTTCAAGTCTTCAAGTCTTCAAGTATGATGATGATGATGATGATGATGAT 6214
5180 CGGAGAGAGATTTTCATGTAATAATTTCTCTGACTTTTGGTTCATCAGTAGACTCGAATCG 5239
6215 CGGAGAGAGATTTTCATGTAATAATTTCTCTGACTTTTGGTTCATCAGTAGACTCGAATCG 6274
5240 TGAGACTATCTCGGTTATGACAGCAGAAATGTCCTTCTTGGAGACAGTAATGAAGTCCC 5299
6275 TGAGACTATCTCGGTTATGACAGCAGAAATGTCCTTCTTGGAGACAGTAATGAAGTCCC 6334
5300 ACCAATAAGAAATCCTTGTATCAGGAAACAACTTCTTGTTCGAACTTTTTCGGTGCC 5359
6335 ACCAATAAGAAATCCTTGTATCAGGAAACAACTTCTTGTTCGAACTTTTTCGGTGCC 6394
5360 TTGAACTATAAATCTAGAGTGGATATCTCGGTAGGAATGGAATGGAATGGAATGGAATG 5419
6395 TTGAACTATAAATCTAGAGTGGATATCTCGGTAGGAATGGAATGGAATGGAATGGAATG 6454
5420 TTCTCGAATCTCAAGAGTATGAGGTTTGTAGGTTTGTAGTATGATGATGATGATGATGATG 5479
6455 TTCTCGAATCTCAAGAGTATGAGGTTTGTAGGTTTGTAGTATGATGATGATGATGATGATG 6514
5480 GTTGTCTTTTCTTCAAAACCAATTCGGAATCCAGAGAAATCAAAAGTGTGTTCTTACTATT 5539
6515 GTTGTCTTTTCTTCAAAACCAATTCGGAATCCAGAGAAATCAAAAGTGTGTTCTTACTATT 6574
5540 GATCAAGCAGTGGCTTGTGAACTGCAATATGATGATGATGATGATGATGATGATGATGAT 5599


```
QY 7760 AACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCACATCGTGCACCCA 7819
|
|
Db 8795 AACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCACATCGTGCACCCA 8854
|
|
QY 7820 ACTGATCTTCAGATCTTTTACTTTTACCTACAGCGTTTCTGGGTGAGCAAAACAGGAGGC 7879
|
|
Db 8855 ACTGATCTTCAGATCTTTTACTTTTACCTACAGCGTTTCTGGGTGAGCAAAACAGGAGGC 8914
|
|
QY 7880 AABATGCGCAAAAGGAATAGGCGCACACGGAATGTTGAATCTCATCTCTTCC 7939
|
|
Db 8915 AAAATGCGCAAAAGGAATAGGCGCACACGGAATGTTGAATCTCATCTCTTCC 8974
|
|
QY 7940 TTTTTCATATATTGAAGCATTTATCAGGGTTATTCTCATGAGCGGATACATATTG 7999
|
|
Db 8975 TTTTTCATATATTGAAGCATTTATCAGGGTTATTCTCATGAGCGGATACATATTG 9034
|
|
QY 8000 AATGATTTAGAAAATAAACAATAGGGTTCGCGCACATTTCCCGAAGAGTCCAC 8059
|
|
Db 9035 AATGATTTAGAAAATAAACAATAGGGTTCGCGCACATTTCCCGAAGAGTCCAC 9094
|
|
QY 8060 CTGACGCTTAAGAAACCATTTATCATGACATTAACCTATAAAATAGCGGTATCACGA 8119
|
|
Db 9095 CTGACGCTTAAGAAACCATTTATCATGACATTAACCTATAAAATAGCGGTATCACGA 9154
|
|
QY 8120 GGCCCTTCGCTTCAAGATTAATTTCTCATGTTGACAGTTATCATCATGAGCTGAC 8179
|
|
Db 9155 GGCCCTTCGCTTCAAGATTAATTTCTCATGTTGACAGTTATCATCATGAGCTGAC 9214
|
|
QY 8180 TCATGTTGTTATTTGAAATAGACGCGAGATCGGGAACACTGAAATAAACAAGTTATT 8239
|
|
Db 9215 TCATGTTGTTATTTGAAATAGACGCGAGATCGGGAACACTGAAATAAACAAGTTATT 9274
|
|
QY 8240 CG.8241
|
|
Db 9275 CG 9276

RESULT 10
AAL60801
ID AAL60801 standard; DNA; 9276 BP.
AC AAL60801;
XX
XX
DT 03-SEP-2003 (first entry)
XX
XX Pichia pastoris vector pPIC9K DNA.
DE Serine protease 17; CVSP17; tumour; cancer; antisense therapy; prostate;
KW breast; cervix; lung; ovary; colon; gene therapy; pPIC9K vector; ds.
XX
XX Pichia pastoris.
FH Key Location/Qualifiers
FT promoter 1..948
|
|
FT primer_bind /tag= a
|
|
FT /note= "5' AOX1 promoter fragment"
|
|
FT 855..875
|
|
FT /tag= v
|
|
FT /bound_moiety= "AOX1 primer"
|
|
FT 949..1218
|
|
FT /tag= b
|
|
FT /note= "Alpha-factor secretion signal"
|
|
FT 1152..1172
|
|
FT /tag= v
|
|
FT /bound_moiety= "Alpha-factor primer"
|
|
FT 1192..1241
|
|
FT /tag= b
|
|
FT /note= "Multiple cloning site"
|
|
FT 1253..1586
|
|
FT /tag= v
|
|
FT /note= "AOX1 transcription termination region"
|
|
FT 1327..1347
|
|
FT /tag= v
```

```
FT primer_bind /bound_moiety= "AOX1 primer"
|
|
FT complement(1980..4514)
|
|
FT /tag= v
|
|
FT /note= "HIS4 ORF"
|
|
FT complement(4928..5743)
|
|
FT /tag= v
|
|
FT /note= "Kanamycin resistance gene"
|
|
FT 6122..6879
|
|
FT /tag= a
|
|
FT /note= "3' AOX1 promoter fragment"
|
|
FT complement(7288..7961)
|
|
FT /tag= v
|
|
FT /note= "ColE1 origin"
|
|
FT complement(8106..8966)
|
|
FT /tag= v
|
|
FT /note= "Ampicillin resistance gene"
|
|
XX WO2003044179-A2.
XX
XX 30-MAY-2003.
XX
XX 20-NOV-2002; 2002WO-US037626.
XX
XX 20-NOV-2001; 2001US-0332015P.
XX
XX (CORV-) CORVAS INT INC.
XX
XX Madison EL, Ong EO;
XX
XX WPI; 2003-449816/42.
XX
XX New substantially purified serine protease 17 polypeptide and encoding
XX nucleic acid, useful for diagnosing and treating tumor conditions and/or
XX cancer, particularly of the breast, cervix, prostate, lung, ovary or
XX colon.
XX
XX Example 2; Page 185-188; 189pp; English.
XX
XX The invention relates to serine protease 17 polypeptide designated CVSP17
XX and its corresponding nucleic acid sequence. The invention also relates
XX to a method using CVSP17 protein to identify compounds that modulate its
XX protease activity. The method is useful for preventing, diagnosing and
XX treating disorders related to the serine protease 17 activity, such as
XX tumour conditions and/or cancer, particularly of the breast, prostate,
XX cervix, lung, ovary or colon. CVSP17 DNA is used in gene therapy and in
XX antisense therapy. The present sequence is Pichia pastoris vector pPIC9K
XX DNA. This sequence is used to illustrate the method of the invention
XX
XX
XX Sequence 9276 BP; 2493 A; 2197 C; 2078 G; 2508 T; 0 U; 0 Other;
XX
XX Query Match 79.6%; Score 6561.6; DB 7; Length 9276;
XX Best Local Similarity 84.7%; Pred. NO. 0;
XX Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;
XX
QY 1 AGATCTAATCAATCCAAAGACGAAAGTTGAAATGAACACCTTTTGGCATCCGACATCCACAG 60
|
|
Db 1 AGATCTAATCAATCCAAAGACGAAAGTTGAAATGAACACCTTTTGGCATCCGACATCCACAG 60
|
|
QY 61 GTCCATTCTCACACATAAGTCCAAACGCAACAGAGGGGATACACTAGCAGCAGACCGT 120
|
|
Db 61 GTCCATTCTCACACATAAGTCCAAACGCAACAGAGGGGATACACTAGCAGCAGACCGT 120
|
|
QY 121 TGCAAACGAGGACCTCCACCTCTTCTCTCAACACCCACTTTTGGCATCCGACATCCACAG 180
|
|
Db 121 TGCAAACGAGGACCTCCACCTCTTCTCTCAACACCCACTTTTGGCATCCGACATCCACAG 180
|
|
QY 181 AGCCAGATTATGGGCTTGATGGAGCTCGCTCATTCCTCAATTCCTTATTAGGCTACTA 240
|
|
Db 181 AGCCAGATTATGGGCTTGATGGAGCTCGCTCATTCCTCAATTCCTTATTAGGCTACTA 240
|
|
QY 241 ACACCATGACTTTATTAGCTGTCTATCCTGGCCCCCTGGCGAGGTTCATGTTGTTTA 300
|
|
Db 241 ACACCATGACTTTATTAGCTGTCTATCCTGGCCCCCTGGCGAGGTTCATGTTGTTTA 300
|
|
```

QY 301 TTTCGAATGCAACAAAGCTCCGCAATTACACCCGAACATCACTCCAGATGAGGGTTTCTG 360
Db 301 TTTCGAATGCAACAAAGCTCCGCAATTACACCCGAACATCACTCCAGATGAGGGTTTCTG 360
QY 361 AGTGTGGGTCAAATAGTTTCATGTTCCCAAAATGGCCCAAACTGACAGTTTAAACGCT 420
Db 361 AGTGTGGGTCAAATAGTTTCATGTTCCCAAAATGGCCCAAACTGACAGTTTAAACGCT 420
QY 421 GTCTTGGAACTTAATATGACAAAAGCGTGATCTCATCCAAAGATGAACTAAAGTTGGTTG 480
Db 421 GTCTTGGAACTTAATATGACAAAAGCGTGATCTCATCCAAAGATGAACTAAAGTTGGTTG 480
QY 481 TTGAAATGCTAACCGCGAGTTGTCACAAAAGAAACTTCCAAAAGTCGCCATACCCTTTGT 540
Db 481 TTGAAATGCTAACCGCGAGTTGTCACAAAAGAAACTTCCAAAAGTCGCCATACCCTTTGT 540
QY 541 CTGTGTTGGTATTGATTCAGCAATGCTCAAAAATAATCTCAATTAATGTTAGCGCAGTCT 600
Db 541 CTGTGTTGGTATTGATTCAGCAATGCTCAAAAATAATCTCAATTAATGTTAGCGCAGTCT 600
QY 601 CTCTATCGCTTCTGNAACCCCGGTGCACTGTGCGAAACGCAAAATGGGAAACACCCGCT 660
Db 601 CTCTATCGCTTCTGNAACCCCGGTGCACTGTGCGAAACGCAAAATGGGAAACACCCGCT 660
QY 661 TTTTGGATGATTATGCTATGCTCCACATTTGATGCTTCCAAAGATTTCTGGTGGGAATACT 720
Db 661 TTTTGGATGATTATGCTATGCTCCACATTTGATGCTTCCAAAGATTTCTGGTGGGAATACT 720
QY 721 GCTGATAGCCTAAAGTTTCATGATCAAAATTTAACTGTTTAAACCCCTACTTGAAGCAAT 780
Db 721 GCTGATAGCCTAAAGTTTCATGATCAAAATTTAACTGTTTAAACCCCTACTTGAAGCAAT 780
QY 781 ATATAACAGAGGAAGTCCCTGCTTAAACCTTTTTTTTATCATCATTTAGCTT 840
Db 781 ATATAACAGAGGAAGTCCCTGCTTAAACCTTTTTTTTATCATCATTTAGCTT 840
QY 841 ACTTTTCATAATTCGACTGGTTTCCAAATGCAAGCTTTTGATTTTAAACGACTTTTAAAGA 900
Db 841 ACTTTTCATAATTCGACTGGTTTCCAAATGCAAGCTTTTGATTTTAAACGACTTTTAAAGA 900
QY 901 CAACTTGAGAGATCAAAAACAACTAATTTTGAAGGATCCAAACGATGAGATTTCCCT 960
Db 901 CAACTTGAGAGATCAAAAACAACTAATTTTGAAGGATCCAAACGATGAGATTTCCCT 960
QY 961 TCAATTTTACTGCAATTTTATTCGAGATCTCCGCAATGCTGCTCCAGTCAACACT 1020
Db 961 TCAATTTTACTGCAATTTTATTCGAGATCTCCGCAATGCTGCTCCAGTCAACACT 1020
QY 1021 ACAACAGAGATGAACCGCAAAATTCGGCTGAAGCTGTCATCGGTTACTCAGATTTA 1080
Db 1021 ACAACAGAGATGAACCGCAAAATTCGGCTGAAGCTGTCATCGGTTACTCAGATTTA 1080
QY 1081 GAAGGGGATTTGATGTTGCTGTTTGGCATTTTCCAAACAGACAAATAACGGGTTATTG 1140
Db 1081 GAAGGGGATTTGATGTTGCTGTTTGGCATTTTCCAAACAGACAAATAACGGGTTATTG 1140
QY 1141 TTTATAAATFACATATTGCCAGCATTTGCTGCTAAAGAAGAGGGGTATCTCTGAGAAA 1200
Db 1141 TTTATAAATFACATATTGCCAGCATTTGCTGCTAAAGAAGAGGGGTATCTCTGAGAAA 1200
QY 1201 AGAGAGGCTGAAGCCCGCAGCCCGCAGCCAGGCCCCAGGAGGAAGAAACATGATCATG 1260
Db 1201 AGAGAGGCTGAAGCTTA ----- 1217
QY 1261 GCCCCCCGGAGAGGATAAATTTGGCTTCCCGGCTGTCACCGCCCCAGCAGTGCACGGAG 1320
Db 1218 ----- 1217
QY 1321 AGAGGTTGCTGTTTGTGATGACAGTGTCCGGGATTTCCCGTGGTGTTCACCCCATGGCC 1380
Db 1218 ----- 1217

QY 1381 ATCGAAGACACTCAAGAAGAAGATGTCCCTTCTAACTAGTGGCTGAGATTTCCCTAGG 1440
Db 1218 ----- CGTAGAATTTCCCTTAGG 1234
QY 1441 CGGCCCGGAATTAATTCGCTTTAGACATGACTGTTCCTCAGTTCAAGTTGGGCACTTAG 1500
Db 1235 CGGCCCGGAATTAATTCGCTTTAGACATGACTGTTCCTCAGTTCAAGTTGGGCACTTAG 1294
QY 1501 AGAAGACCGCTCTTCTAGATTTCTAATCAAGAGGATGTCAGATGCCATTTCCCTGAGAG 1560
Db 1295 AGAAGACCGCTCTTCTAGATTTCTAATCAAGAGGATGTCAGATGCCATTTCCCTGAGAG 1354
QY 1561 ATCGAGGCTTCATTTTGTATCTTTTATTTGTAACCTATATAGTATAGGATTTTTTTT 1620
Db 1355 ATCGAGGCTTCATTTTGTATCTTTTATTTGTAACCTATATAGTATAGGATTTTTTTT 1414
QY 1621 GTCAATTTGTTTCTTCTCTAGAGCTTCTCTCTGATCAGCCTATCTCCAGCTGATGAA 1680
Db 1415 GTCAATTTGTTTCTTCTCTAGAGCTTCTCTCTGATCAGCCTATCTCCAGCTGATGAA 1474
QY 1681 TATCTTGTGTAGGGTTTGGGAAATCATTCGAGTTTGTATGTTTCTTGGTATTTCC 1740
Db 1475 TATCTTGTGTAGGGTTTGGGAAATCATTCGAGTTTGTATGTTTCTTGGTATTTCC 1534
QY 1741 ACTCTCTTTCAGAGTACAGAATTAAGTGAAGTTCGTTTGTGCAAGCTTATCGATAA 1800
Db 1535 ACTCTCTTTCAGAGTACAGAATTAAGTGAAGTTCGTTTGTGCAAGCTTATCGATAA 1594
QY 1801 GCTTTAATGGGTAGTTTATCAGTTAAATTCGCTAAACGAGTCAGGACCGTGTATGAA 1860
Db 1595 GCTTTAATGGGTAGTTTATCAGTTAAATTCGCTAAACGAGTCAGGACCGTGTATGAA 1654
QY 1861 ATCTAACAATCGCTCATCTCGGACCGCTCACCTCGATGCTGTAGGCAATAG 1920
Db 1655 ATCTAACAATCGCTCATCTCGGACCGCTCACCTCGATGCTGTAGGCAATAG 1714
QY 1921 CTTGGTTATGCGGTAAGTCCCGGCTCTTGGGATATCGTCCATTCCGACAGCATCGC 1980
Db 1715 CTTGGTTATGCGGTAAGTCCCGGCTCTTGGGATATCGTCCATTCCGACAGCATCGC 1774
QY 1981 CAGTCACTATGCGGCTGCTGCTAGCGCTATATGCTGATGCAATTTCTATGCGCACCGCT 2040
Db 1775 CAGTCACTATGCGGCTGCTGCTAGCGCTATATGCTGATGCAATTTCTATGCGCACCGCT 1834
QY 2041 TCTCGAGCACTGCTCGACCGCTTTGGCGCGCGCCAGTCTGCTCGCTTCTGCTACTTGG 2100
Db 1835 TCTCGAGCACTGCTCGACCGCTTTGGCGCGCGCCAGTCTGCTCGCTTCTGCTACTTGG 1894
QY 2101 AGCCACTATCGACTACCGGATCATGCGGACCAACCCGCTCTGTGATCTATCGAATCTA 2160
Db 1895 AGCCACTATCGACTACCGGATCATGCGGACCAACCCGCTCTGTGATCTATCGAATCTA 1954
QY 2161 AATGTAAGTTAAATCTCTAAATTAATTAAGTCCCAAGTTTCTCCATACGAACCTTAA 2220
Db 1955 AATGTAAGTTAAATCTCTAAATTAATTAAGTCCCAAGTTTCTCCATACGAACCTTAA 2014
QY 2221 CAGCATTTGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCCATCTGTTGGCCAA 2280
Db 2015 CAGCATTTGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCCATCTGTTGGCCAA 2074
QY 2281 TATGTTTTCAGTCCCTCAGGAGTTTACGCTTTGTGAAGTGAATGAACCTTCTGGAAGTTGCGAG 2340
Db 2075 TATGTTTTCAGTCCCTCAGGAGTTTACGCTTTGTGAAGTGAATGAACCTTCTGGAAGTTGCGAG 2134
QY 2341 TGTAACTCCGCTGATTTGACGGGCTATCCGTTACGTTGGCAAGTGTGGTGGTACCGG 2400
Db 2135 TGTAACTCCGCTGATTTGACGGGCTATCCGTTACGTTGGCAAGTGTGGTGGTACCGG 2194
QY 2401 AGAGGTAATCTCCCAACTCTCTGGAGAGTAGGACCAACAAACAGATCCAGAGCTGTT 2460
Db 2195 AGAGGTAATCTCCCAACTCTCTGGAGAGTAGGACCAACAAACAGATCCAGAGCTGTT 2254
QY 2461 GTACTTTGATCAACATAGAAGAGCATTTCTCGATTTTTCAGGATCAAGTGTTCAGGAGCT 2520

Qy	4681	CTGCAACACTTGGCTATGTCAGGTAGCAAGGAAATGTACATCTTGAAGTCGGACAGTGAG	4740	Qy	4829	-----	4828
Dd	4475	CTGCAACACTTGGCTATGTCAGGTAGCAAGGAAATGTACATCTTGAAGTCGGACAGTGAG	4534	Dd	5555	ATTGGCAACGCTACCTTTTGCCATGTTTCAGAAACAACCTCTGGCGCATCGGCTTCCCATTA	5614
Qy	4741	TGTAGTCTTGTGAGAAATCTCTGAAGCCGTATTTTATATACGTAGTCACTCATCAGGAGA	4800	Qy	4829	-----	4828
Dd	4535	TGTAGTCTTGTGAGAAATCTCTGAAGCCGTATTTTATATACGTAGTCACTCATCAGGAGA	4594	Dd	5615	CAATCGATAGATTGTGCACTGATTGCCGACATTAATCGCAGGCCCATTTATACCCATA	5674
Qy	4801	TCCTCTACGCGGACGATCTGTGCGCA-----	4828	Qy	4829	-----	4828
Dd	4595	TCCTCTACGCGGACGATCTGTGCGCA-----	4828	Dd	5675	TAAATCAGCATCCATGTTGGAATTTAAATCGCGGCTCGAGCAAGACGTTTCCGCTTGAAT	5734
Qy	4829	-----	4828	Qy	4829	-----	4828
Dd	4655	CTCGTGAAGAAGTGTGCTGACTCATACCAGGCTGAATCGCCCATCATCCAGCCAGA	4714	Dd	5735	ATGGCTCATACACCCCTTGTATTAAGTATTAAGCAGACAGTTTATTTGTTTCATGA	5794
Qy	4829	-----	4828	Qy	4829	-----	4828
Dd	4715	AAGTAGGAGGCCACGGTTGATGAGAGCTTTGTTGTAGGTGGAACAGTTGGTGAATTTGA	4774	Dd	5795	TGATATATTTTATCTTGTGCAATGTAAACATCAGAGATTTTGAGACACAAACGTTGGCTTC	5854
Qy	4829	-----	4828	Qy	4829	-----	4828
Dd	4775	ACTTTTGCTTTGCCACGGAACGGTCTGCTGCTGCGGAAGATGCGTGTATCTGATCTGATCCTCA	4834	Dd	5855	CCCCCCCCCTGACAGTCCGATCAACCGGCCCAAGTGGCTTGTGGCGCCTATAT	5914
Qy	4829	-----	4828	Qy	4880	CGCCGACATCACCGATGGGGAAGATCGGCTCGCCACTTCGGGCTCATGAGGCTTGT	4939
Dd	4835	ACTCAGCAAAAGTTCGATTTATTCACAAAGCCGCGTCCGTCAGGTCAAGTCAATGCT	4894	Dd	5915	CGCCGACATCACCGATGGGGAAGATCGGCTCGCCACTTCGGGCTCATGAGGCTTGT	5974
Qy	4829	-----	4828	Qy	4940	CGCGTGGGTATGTCGACAGGCCCTCGTGGCGGGGACTGTTGGGCGGCATCTCCTTGA	4999
Dd	4895	CTGCGAGTTTACAAACCAATTAACCAATCTGATTAGAAAACATCATCGACATCAATG	4954	Dd	5975	CGCGTGGGTATGTCGACAGGCCCTCGTGGCGGGGACTGTTGGGCGGCATCTCCTTGA	6034
Qy	4829	-----	4828	Qy	5000	TGACACATTCCTTGGCGCGGCTGCTCAACGGCCCTCAACCTACTACTGGGCTTCTCT	5059
Dd	4955	AACTGCAATTTATTCATATCAGGATTAATCAATACCATATTTTGAAAAAGCCGTTCTG	5014	Dd	6035	TGACACATTCCTTGGCGCGGCTGCTCAACGGCCCTCAACCTACTACTGGGCTTCTCT	6094
Qy	4829	-----	4828	Qy	5060	AATGCGAGGTGTCATGAAGGAGGAGCTCGAGTATCTATATTTGGAAGTATGGAATGTT	5119
Dd	5015	TAATGAAGGAGAAACTCACGAGGAGTTCATAGGATGGCAAGATCCTGGTATCGGTC	5074	Dd	6095	AATGCGAGGTGTCATGAAGGAGGAGCTCGAGTATCTATATTTGGAAGTATGGAATGTT	6154
Qy	4829	-----	4828	Qy	5120	GATACCGCATTCCTCAGTGTCTTGGTCTCTCTATCAGATTTGCCCACTAAAGCAAC	5179
Dd	5075	TGCGATTCCGACTCGTCCACATCAATCAACCTATTATTTCCCTCGTCAAAAATAAG	5134	Dd	6155	GATACCGCATTCCTCAGTGTCTTGGTCTCTCTATCAGATTTGCCCACTAAAGCAAC	6214
Qy	4829	-----	4828	Qy	5180	CGAGGAGGAGATTTCTATGTTAAATTTCTGACTTTTGGTCTCATCTAGTAGACTCGAACTG	5239
Dd	5135	GTATCAAGTGAGAAATCACCATGAGTACGAGTCACTGAATCCGGTGAGAAATGCAAAAGCTT	5194	Dd	6215	CGAGGAGGAGATTTCTATGTTAAATTTCTGACTTTTGGTCTCATCTAGTAGACTCGAACTG	6274
Qy	4829	-----	4828	Qy	5240	TGAGACTATCTCGGTTATGACAGCAAAATGTCTTCTTGGGAGACAGTAAATGAACTCC	5299
Dd	5195	ATGCAATTTCTTCCAGACTTGTTCACAGGCCAGCCATACGCTCGTCAATCAAAATCACT	5254	Dd	6275	TGAGACTATCTCGGTTATGACAGCAAAATGTCTTCTTGGGAGACAGTAAATGAACTCC	6334
Qy	4829	-----	4828	Qy	5300	ACCAATAAGAAATCCTTGTATCAGGAACAACTCTTGTTCGAACTTTTTCGGTGCC	5359
Dd	5255	CGCATCAACAAACCGTTATTCATTCTGATTGCGCTGAGCGAGACGAAATACCGGATC	5314	Dd	6335	ACCAATAAGAAATCCTTGTATCAGGAACAACTCTTGTTCGAACTTTTTCGGTGCC	6394
Qy	4829	-----	4828	Qy	5360	TTGAACTTATATAATGTAGAGTGGATATGTCGGTAGGAAATGGAGCGGGCAAAATGCTTACC	5419
Dd	5315	GCTGTTAAAGGACAAATTAACAAAGGAATCGAATGCAACCGCGGCGCAGAACACTGCCAG	5374	Dd	6395	TTGAACTTATATAATGTAGAGTGGATATGTCGGTAGGAAATGGAGCGGGCAAAATGCTTACC	6454
Qy	4829	-----	4828	Qy	5420	TTCTGACCTTCAAGAGGTATGAGGTTTGTAGTACTGATGCCAATTCAGTCACTTACGAGCAAC	5479
Dd	5375	CGCATCAACAAATTTTACCTGAATCAGGATATCTTCTTAATACCTGGAATGCTGTTTT	5434	Dd	6455	TTCTGACCTTCAAGAGGTATGAGGTTTGTAGTACTGATGCCAATTCAGTCACTTACGAGCAAC	6514
Qy	4829	-----	4828	Qy	5480	GTTGCTATTTTCGTTCAAAACCATTCGGAATCCAGAGAAATCAAAAGTTGTTTGTCTACTATT	5539
Dd	5435	CCCGGGATCGCAGTGGTGAATACCAAGCATCATCAGGAGTACGGATAAAATGTTGAT	5494	Dd	6515	GTTGCTATTTTCGTTCAAAACCATTCGGAATCCAGAGAAATCAAAAGTTGTTTGTCTACTATT	6574
Qy	4829	-----	4828	Qy	5540	GATCAAGCCAGTGGCTTGTGAACTGACAAATAGTGTGCTGCTGTTTGGAGTCACTTT	5599
Dd	5495	GGTCGAGAGGCAATAATTCGTCAGCCAGTTTGTAGTCTGACCACTCATCTGTAAATC	5554	Dd	6575	GATCAAGCCAGTGGCTTGTGAACTGACAAATAGTGTGCTGCTGTTTGGAGTCACTTT	6634

Db 6635 TGTATGAATAAATCTAGTCTTTGATCTAAATAAATCTTTGACGAGCAAGGCGATAAATAACC 6694
Qy 5660 CAATCTAAATCTCTTTAAACGTTAAAGGACAAGTATGTCGCTGTATTAACCC 5719
Db 6695 CAATCTAAATCTCTTTAAACGTTAAAGGACAAGTATGTCGCTGTATTAACCC 6754
Qy 5720 AAATCAGCTCTAGTCTGATCCTCATCAACCTTTAGGGGCACTATCTTTGTTTATAGAGAAAT 5779
Db 6755 AAATCAGCTCTAGTCTGATCCTCATCAACCTTTAGGGGCACTATCTTTGTTTATAGAGAAAT 6814
Qy 5780 TTGGGAGATCGGATATCGAGAAAAGGTAGCTGTATTTAAAGCTGAATTTATCTCAA 5839
Db 6815 TTGGGAGATCGGATATCGAGAAAAGGTAGCTGTATTTAAAGCTGAATTTATCTCAA 6874
Qy 5840 GATCTCTCCCTCGGCGTTTCGGTGTATGACGGTGAACCTCTGACACATGCGAGCTCCG 5899
Db 6875 GATCTCTCCCTCGGCGTTTCGGTGTATGACGGTGAACCTCTGACACATGCGAGCTCCG 6934
Qy 5900 GAGAGGTCACAGCTTGCTGTAAAGCGGATGCGGGAGCAGCAAGCCCGTCTAGGGCGCG 5959
Db 6935 GAGAGGTCACAGCTTGCTGTAAAGCGGATGCGGGAGCAGCAAGCCCGTCTAGGGCGCG 6994
Qy 5960 TCAGCGGTGTGGCGGTGTCGGGCGCAGCATGACCCAGTCACGTCAGCTAGCGAGCGCG 6019
Db 6995 TCAGCGGTGTGGCGGTGTCGGGCGCAGCATGACCCAGTCACGTCAGCTAGCGAGCGCG 7054
Qy 6020 GTGTACTGCTTAATCATCGGCATCAGAGCAGATGTACTGAGAGTGCACCATATGC 6079
Db 7055 GTGTACTGCTTAATCATCGGCATCAGAGCAGATGTACTGAGAGTGCACCATATGC 7114
Qy 6080 GGTGTGAATACCGCAGATGCGTAGGAGAAAATACCGCATCAGCGGCTCTCCGCTT 6139
Db 7115 GGTGTGAATACCGCAGATGCGTAGGAGAAAATACCGCATCAGCGGCTCTCCGCTT 7174
Qy 6140 CCTCGCTCACTGACTCGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCACT 6199
Db 7175 CCTCGCTCACTGACTCGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCACT 7234
Qy 6200 CAAGGCGGTAAATACGTTATTCACAGAAATCAGGGGATTAACGAGGAAAGAAATGTGAG 6259
Db 7235 CAAGGCGGTAAATACGTTATTCACAGAAATCAGGGGATTAACGAGGAAAGAAATGTGAG 7294
Qy 6260 CAAGAGGCGCAGAAAGCCAGGAAACCGTAAAGAGGCGGCTGCTGGCGTTTTCATA 6319
Db 7295 CAAGAGGCGCAGAAAGCCAGGAAACCGTAAAGAGGCGGCTGCTGGCGTTTTCATA 7354
Qy 6320 GGCTCCGCCCCCTGAGCAGATCAAAAATCGACGCTAAGTCAGAGTGGCGAAACC 6379
Db 7355 GGCTCCGCCCCCTGAGCAGATCAAAAATCGACGCTAAGTCAGAGTGGCGAAACC 7414
Qy 6380 CGACAGGACTAAAGATACAGAGCTTTCCCTCGGAAGCTCCCTCGGCTCTCCTG 6439
Db 7415 CGACAGGACTAAAGATACAGAGCTTTCCCTCGGAAGCTCCCTCGGCTCTCCTG 7474
Qy 6440 TTCCGACCTCGGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGC 6499
Db 7475 TTCCGACCTCGGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGC 7534
Qy 6500 TTTCTCATAGCTCAGCTGTAGTATCTCAGTTCGGTGTAGTTCGCTCCAGCTGG 6559
Db 7535 TTTCTCAATGTCTAGCTGTAGTATCTCAGTTCGGTGTAGTTCGCTCCAGCTGG 7594
Qy 6560 GCTGTGTGACGAAACCCCGCTTCAGCGGACCGCTGCGGCTTATCCGTTATCTGTC 6619
Db 7595 GCTGTGTGACGAAACCCCGCTTCAGCGGACCGCTGCGGCTTATCCGTTATCTGTC 7654
Qy 6620 TTGAGTCCAAACCGGTATAGACAGACTTATGCCACTGGCAGGACCTGGTAAACAGGA 6679
Db 7655 TTGAGTCCAAACCGGTATAGACAGACTTATGCCACTGGCAGGACCTGGTAAACAGGA 7714
Qy 6680 TTAGCAGAGCGAGGTATGTAGCGGTGTACAGAGTTCCTGAAAGTGGTGGCTTAACCTACG 6739

Db 7715 TTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCCTGAAAGTGGTGGCTTAACCTACG 7774
Qy 6740 GCTACACTAGAGGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTATACCTTCGGAA 6799
Db 7775 GCTACACTAGAGGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTATACCTTCGGAA 7834
Qy 6800 AAAGAGTTGGTAGCTCTTTGATCCGSCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTG 6859
Db 7835 AAAGAGTTGGTAGCTCTTTGATCCGSCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTG 7894
Qy 6860 TTTGCAAGCAGCAGATTAACGCGCAGAAAAAAGAGATCTCAAGAAAGATCCTTTGATCTTTT 6919
Db 7895 TTTGCAAGCAGCAGATTAACGCGCAGAAAAAAGAGATCTCAAGAAAGATCCTTTGATCTTTT 7954
Qy 6920 CTACGGGCTCTGACGCTCAGTGGAAACGAAACTCAAGTAAAGGATTTTGGTCTATGAGAT 6979
Db 7955 CTACGGGCTCTGACGCTCAGTGGAAACGAAACTCAAGTAAAGGATTTTGGTCTATGAGAT 8014
Qy 6980 TATCAAAAGAGATCTTCACTAGATCCTTTTAAATTAATAAATGAAGTTTAAATCAATCT 7039
Db 8015 TATCAAAAGAGATCTTCACTAGATCCTTTTAAATTAATAAATGAAGTTTAAATCAATCT 8074
Qy 7040 AAAGTATATATGAGTAAACTTTGGTCTGACAGTTTACCAATGCTTAAATCAGTCAGGACCTA 7099
Db 8075 AAAGTATATATGAGTAAACTTTGGTCTGACAGTTTACCAATGCTTAAATCAGTCAGGACCTA 8134
Qy 7100 TCTCAGCGATCTGTCTATTTGTTTCATCCATAGTTGCTGCTGCTCCCGTCTGTAGATAA 7159
Db 8135 TCTCAGCGATCTGTCTATTTGTTTCATCCATAGTTGCTGCTGCTCCCGTCTGTAGATAA 8194
Qy 7160 CTACCATACGGAGGCGCTTACCATCTGCCCCCAGTGTGCAATGATATCCGCGAGACCCAC 7219
Db 8195 CTACCATACGGAGGCGCTTACCATCTGCCCCCAGTGTGCAATGATATCCGCGAGACCCAC 8254
Qy 7220 GTCACCGGCTCCAGATTTATCAGCAATTAACACGAGCAGCCGGAAGGCGCAGGCGCAGAA 7279
Db 8255 GTCACCGGCTCCAGATTTATCAGCAATTAACACGAGCAGCCGGAAGGCGCAGGCGCAGAA 8314
Qy 7280 GGTGCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGCGGAAGCTAGAG 7339
Db 8315 GGTGCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGCGGAAGCTAGAG 8374
Qy 7340 TAAAGTGTGCGCACTAATAGTTGCGCAAGCTTTGTCGCAATGCTGCTCAGGCAATCGTGG 7399
Db 8375 TAAAGTGTGCGCACTAATAGTTTGGCAACCTTTGTTGCCATGCTGCTCAGGCAATCGTGG 8434
Qy 7400 TGTGAGCTCCGCTGCTGATGCTTCAATTCAGCTCCGCTTCCCAAGATCAAGGCGAG 7459
Db 8435 TGTGAGCTCCGCTGCTGATGCTTCAATTCAGCTCCGCTTCCCAAGATCAAGGCGAG 8494
Qy 7460 TTACATGATCCCCCATGTTGTCGCAAAAAGCGGTTAGCTCTCGGCTCTCCGATCGTTG 7519
Db 8495 TTACATGATCCCCCATGTTGTCGCAAAAAGCGGTTAGCTCTCGGCTCTCCGATCGTTG 8554
Qy 7520 TCAGAAAGTGTGCGCGCAGTGTATCACTCATGTTTATGCGCAGCTGCGATATCTC 7579
Db 8555 TCAGAAAGTGTGCGCGCAGTGTATCACTCATGTTTATGCGCAGCTGCGATATCTC 8614
Qy 7580 TTACTGTCAATGCGCATCCGTAAGATGCTTTTGTGACCTGGTGAAGTCTCAACCAAGTCAAT 7639
Db 8615 TTACTGTCAATGCGCATCCGTAAGATGCTTTTGTGACCTGGTGAAGTCTCAACCAAGTCAAT 8674
Qy 7640 TGTGAAATAGTGTATGCGGCGACCGAGTTGCTTTCGCGGCTGCTCAACGAGGATATA 7699
Db 8675 TGTGAAATAGTGTATGCGGCGACCGAGTTGCTTTCGCGGCTGCTCAACGAGGATATA 8734
Qy 7700 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATTTGGAAGAGCTTTCTTCGGGGCGAA 7759
Db 8735 CCGCGCCACATAGCAGAACTTTAAAGTGTCTCATTTGGAAGAGCTTTCTTCGGGGCGAA 8794
Qy 7760 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATTAACCCACTCGTGCACCCA 7819
Db 8795 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATTAACCCACTCGTGCACCCA 8854


```

FT FT /tag= f
FT /bound_moiety= "AOX1 primer"
FT complement(1980. .4514)
FT /tag= h
FT /note= "HIS4 ORF"
FT misc_feature complement(4928. .5743)
FT /tag= i
FT /note= "kanamycin resistance gene"
FT 6122. .6879
FT /tag= j
FT /note= "AOX1 fragment"
FT misc_feature complement(7288. .7961)
FT /tag= k
FT /note= "ColE1 origin"
FT misc_feature complement(8106. .8966)
FT /tag= l
FT /note= "Ampicillin resistance gene"
XX WO200277267-A2.
XX
XX
XX 03-OCT-2002.
XX
XX 27-MAR-2002; 2002WO-US009611.
XX
XX 27-MAR-2001; 2001US-0279228P.
XX
XX 15-MAY-2001; 2001US-0291501P.
XX
XX (CORV-) CORVAS INT INC.
XX
XX Madison EL, Ong EO;
XX
XX WPI; 2003-018940/01.
XX
XX New substantially purified single or two-chain type II membrane-type
XX serine protease 9 (MTSP9) polypeptide, useful for monitoring tumor
XX progression, inhibiting tumor initiation, or treating a malignant or pre-
XX malignant condition.
XX
XX Example 2; Page 191-193; 199pp; English.
XX
XX The invention relates to type II membrane-type serine protease 9 (MTSP9)
XX polypeptides and polynucleotides. MTSP9 belongs to type II transmembrane
XX serine protease (TSP) family. Sequences of the invention and their
XX antibodies are useful for diagnosing, treating or preventing neoplastic
XX disease in mammals. They are useful for monitoring tumour progression,
XX inhibiting tumour initiation, growth or progression or treating malignant
XX or pre-malignant conditions. Transgenic animals of the invention are
XX useful in animal models of tumour initiation, growth and/or progression
XX models. The invention is also useful in gene therapy. The present
XX sequence is pPIC9K vector DNA containing human MTSP9 sequence
XX
XX Sequence 9276 BP; 2493 A; 2197 C; 2078 G; 2508 T; 0 U; 0 Other;
XX
XX Query Match 79.6%; Score 6561.6; DB 7; Length 9276;
XX Best Local Similarity 84.7%; Pred. No. 0;
XX Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;
XX
XX 1 AGATCTTAACATCCAAAGACGAAAGGTTGATGAACACCTTTTGCCATCCGACATCCACAG 60
XX |||||
XX 1 AGATCTTAACATCCAAAGACGAAAGGTTGATGAACACCTTTTGCCATCCGACATCCACAG 60
XX
XX 61 GTCCATTCTCACATCAAGTGCCAAACGCAACAGAGGGGATACACTAGCAGCAGACCGT 120
XX |||||
XX 61 GTCCATTCTCACATCAAGTGCCAAACGCAACAGAGGGGATACACTAGCAGCAGACCGT 120
XX
XX 121 TGCAAAACGACGACCTCCACTCTCTCTTCCTCAACACCCACTTTTGCCATCGAAAACC 180
XX |||||
XX 121 TGCAAAACGACGACCTCCACTCTCTCTTCCTCAACACCCACTTTTGCCATCGAAAACC 180
XX
XX 181 AGCCAGTATTGGGCTTCGATTGGAGCTCGCTCATTCCAATTCTTCTATTAGGCTACTA 240
XX |||||
XX 181 AGCCAGTATTGGGCTTCGATTGGAGCTCGCTCATTCCAATTCTTCTATTAGGCTACTA 240

```

QY 241 ACACCATGACCTTTATTAGCCTGTCTATCCCTGGCCCCCTCGCGAGGTTTCATGTTTGTGTTA 300
Db 241 ACACCATGACCTTTATTAGCCTGTCTATCCCTGGCCCCCTCGCGAGGTTTCATGTTTGTGTTA 300
QY 301 TTTCCGAATGCAACAGCTCGCATTAACACCCGAAACATCACTCCAGATGAGGCTTTCTG 360
Db 301 TTTCCGAATGCAACAGCTCGCATTAACACCCGAAACATCACTCCAGATGAGGCTTTCTG 360
QY 361 AGTGGGGTCAATAGTATTTTCATGTTTCCCAATGGCCCAAACTGACAGTTTAAAGCT 420
Db 361 AGTGGGGTCAATAGTATTTTCATGTTTCCCAATGGCCCAAACTGACAGTTTAAAGCT 420
QY 421 GTCTTGGAACTTAATATGACAAAAGCGTGATCTCATCCAAAGATGAACCTAAGTTTGGTTCG 480
Db 421 GTCTTGGAACTTAATATGACAAAAGCGTGATCTCATCCAAAGATGAACCTAAGTTTGGTTCG 480
QY 481 TTGAATGCTAACCGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCCGCAATACCGTTTGT 540
Db 481 TTGAATGCTAACCGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCCGCAATACCGTTTGT 540
QY 541 CTCTGTTGGTATGATGACGAATGCTCAAAAATAATCTCAATTAATGCTTAGCGCAGTCT 600
Db 541 CTCTGTTGGTATGATGACGAATGCTCAAAAATAATCTCAATTAATGCTTAGCGCAGTCT 600
QY 601 CTCTATCGCTTCTGAACCCCGGTGCACTGTGCGGAAACGCAAAATGGGGAAAACCCCGCT 660
Db 601 CTCTATCGCTTCTGAACCCCGGTGCACTGTGCGGAAACGCAAAATGGGGAAAACCCCGCT 660
QY 661 TTTTGGATGATTAATGCAATGCTCCACATGCTATGCTTCCAGANTCTGCTGGGAATACT 720
Db 661 TTTTGGATGATTAATGCAATGCTCCACATGCTATGCTTCCAGANTCTGCTGGGAATACT 720
QY 721 GCTGATAGCCTTAAGCTTCATGATCAAAATTAAGTCTTAACCCCTACTTGACAGCAAT 780
Db 721 GCTGATAGCCTTAAGCTTCATGATCAAAATTAAGTCTTAACCCCTACTTGACAGCAAT 780
QY 781 ATATAACAGAAAGTGGCTGCTTAAACCTTTTATCATCATTAATAGCTT 840
Db 781 ATATAACAGAAAGTGGCTGCTTAAACCTTTTATCATCATTAATAGCTT 840
QY 841 ACTTTCAATTTGCGACTGTTTCAATGACAGCTTTTGAATTAACGACTTTTAAACGA 900
Db 841 ACTTTCAATTTGCGACTGTTTCAATGACAGCTTTTGAATTAACGACTTTTAAACGA 900
QY 901 CAACCTTGAGAGATCAAAAACAACTAATTTTGAAGAGTCCAAAGATGAGATTCTCT 960
Db 901 CAACCTTGAGAGATCAAAAACAACTAATTTTGAAGAGTCCAAAGATGAGATTCTCT 960
QY 961 TCAATTTTACTGAGTTTATTCGAGCATCTCCGCAATAGCTGCTCCAGTCAACACT 1020
Db 961 TCAATTTTACTGAGTTTATTCGAGCATCTCCGCAATAGCTGCTCCAGTCAACACT 1020
QY 1021 ACAACAGAGATGAACCGCAAAATTCGGCTGAAGCTGTCAATCGGTTACTCAGATTTA 1080
Db 1021 ACAACAGAGATGAACCGCAAAATTCGGCTGAAGCTGTCAATCGGTTACTCAGATTTA 1080
QY 1081 GAAGGGATTTGATGTTGCTGTTTCCATTTTCCAAAGAGCAAAATAACGGGTTATTG 1140
Db 1081 GAAGGGATTTGATGTTGCTGTTTCCATTTTCCAAAGAGCAAAATAACGGGTTATTG 1140
QY 1141 TTTATAATACTACTATTGCGCATTTGCTGCTAAAGAGAGGGGTATCTCTCGAGAAA 1200
Db 1141 TTTATAATACTACTATTGCGCATTTGCTGCTAAAGAGAGGGGTATCTCTCGAGAAA 1200
QY 1201 AGAGAGGCTGAAGCCAGGCCAGGCCAGGCCAGGCCAGGAGAAACATGATCATG 1260
Db 1201 AGAGAGGCTGAAGCTA----- 1217
QY 1261 GCCCCCGGGAGAGGATAAATTTGGCTTCCCGGTTGTCACCGCCAGCAGTGCACGGAG 1320
Db 1218 ----- 1217
QY 1321 AGAGGTTGCTGTTTGTATGACAGTGTCCGGGGATTCGGTGGTCTTCCACCCCATGGCC 1380

Db 1218 ----- 1217
QY 1381 ATCGAGAACACTCAAGAAAGAAATGTCCTTTCTAACTAGTGGCGTAGAATTCCTAGGG 1440
Db 1218 -----CGTAGAATTCCTAGGG 1234
QY 1441 CGGCGCGAATTAATTCGCTTTAGACATGACTGTTCTCAGTTCAGTTGGCAGCTTTACG 1500
Db 1235 CGGCGCGAATTAATTCGCTTTAGACATGACTGTTCTCAGTTCAGTTGGCAGCTTTACG 1234
QY 1501 AGAAGACCGGCTCTGCTAGATTCTTAATCAAGAGATGTCAGAAATGCCATTTGCCGTGAG 1560
Db 1295 AGAAGACCGGCTCTGCTAGATTCTTAATCAAGAGATGTCAGAAATGCCATTTGCCGTGAG 1354
QY 1561 ATGCAAGGCTTCATTTTGTGATACCTTTTATTTGTAACCTATATAGTATAGATTTTTTT 1620
Db 1355 ATGCAAGGCTTCATTTTGTGATACCTTTTATTTGTAACCTATATAGTATAGATTTTTTT 1414
QY 1621 GTCAATTTTGTCTTCTCGTAGAGCTTGTCTCGATCAGCCTATCTCGAGCTGATGAA 1680
Db 1415 GTCAATTTTGTCTTCTCGTAGAGCTTGTCTCGATCAGCCTATCTCGAGCTGATGAA 1474
QY 1681 TATCTGTGTAGGGTTTGGGAAATCATTTCGAGTTTGATGTTTCTTGGTATTTCCC 1740
Db 1475 TATCTGTGTAGGGTTTGGGAAATCATTTCGAGTTTGATGTTTCTTGGTATTTCCC 1534
QY 1741 ACTCCTCTTCAGAGTACAGAAAGATTAAAGTGAGAAAGTTTCGTTTGTGCAAGCTTATCGATAA 1800
Db 1535 ACTCCTCTTCAGAGTACAGAAAGATTAAAGTGAGAAAGTTTCGTTTGTGCAAGCTTATCGATAA 1594
QY 1801 GCTTTAATGCGGTAGTTTATCAGAGTTAAATGCTTAAGCGAGTTCAGGACCGTGTATGAA 1860
Db 1595 GCTTTAATGCGGTAGTTTATCAGAGTTAAATGCTTAAGCGAGTTCAGGACCGTGTATGAA 1654
QY 1861 ATCTAAACATGCGCTCATCGTTCCTCGGACCGCTCACCCCTGGATGCTGTAGGCATAGG 1920
Db 1655 ATCTAAACATGCGCTCATCGTTCCTCGGACCGCTCACCCCTGGATGCTGTAGGCATAGG 1714
QY 1921 CTTGTTTATGCGGTACTGCGGGGCTCTTCGGGGATATCGTCCATTCGACAGCATCGC 1980
Db 1715 CTTGTTTATGCGGTACTGCGGGGCTCTTCGGGGATATCGTCCATTCGACAGCATCGC 1774
QY 1981 CAGTCACATGCGGTGCTAGCGTATATGCGTGAATGCAATTTCTATCGCACCCGT 2040
Db 1775 CAGTCACATGCGGTGCTAGCGTATATGCGTGAATGCAATTTCTATGCGCACCCGT 1834
QY 2041 TCTCGGAGCATGTCCGACCGCTTTGGCCGCGCCAGTCTGCTGCTTCGCTTACTTGG 2100
Db 1835 TCTCGGAGCATGTCCGACCGCTTTGGCCGCGCCAGTCTGCTGCTTCGCTTACTTGG 1894
QY 2101 AGCCATATCGACTACCGGATCAGTGGGACCAACCCGCTCTGTTGGATCTATCGAATCTA 2160
Db 1895 AGCCATATCGACTACCGGATCAGTGGGACCAACCCGCTCTGTTGGATCTATCGAATCTA 1954
QY 2161 AATGTAAGTTAAATCTCTAAATTAATTAAGTCCAGTTTCTCCATACGAAACCTTAA 2220
Db 1955 AATGTAAGTTAAATCTCTAAATTAATTAAGTCCAGTTTCTCCATACGAAACCTTAA 2014
QY 2221 CAGCATTCGGGTAGCATCTAGACTTCAACAGAGCCAGATCCATCATCTCTTGGCCAA 2280
Db 2015 CAGCATTCGGGTAGCATCTAGACTTCAACAGAGCCAGATCCATCATCTCTTGGCCAA 2074
QY 2281 TATGTTTCACTCCCTCAGGATTAGCTTCTGTGAAGTGAATCTTCTGGAAGTTGCGAG 2340
Db 2075 TATGTTTCACTCCCTCAGGATTAGCTTCTGTGAAGTGAATCTTCTGGAAGTTGCGAG 2134
QY 2341 TGTAACTCCGCTGTATTTGACGGGCATATCGTACGTTGGCAAGTGTGTTGTACCGG 2400
Db 2135 TGTAACTCCGCTGTATTTGACGGGCATATCGTACGTTGGCAAGTGTGTTGTACCGG 2194
QY 2401 AGGAGTAAATCTCCAACTCTCTCGAGAGTAGGACCAACAAACACAGATCCAGCGTGT 2460

QY	3541	TAGGTTGTTCCCTGGGTGTAATCCTGGCTTCGGCATCTCCTTTCTCTTAGTGACCTTAGGG	3600
DB	3335	TAGGTTGTTCCCTGGGTGTAATCCTGGCTTCGGCATCTCCTTTCTCTTAGTGACCTTAGGG	3394
QY	3601	ACTTCATATCCAGGTTTCTCTCCACCTCGTCCAAACGTCAACGGTACTTGGCACATCTAA	3660
DB	3395	ACTTCATATCCAGGTTTCTCTCCACCTCGTCCAAACGTCAACGGTACTTGGCACATCTAA	3454
QY	3661	CTAATCCAAATAAATAAAGTCAGACACANTCCACGGCTATATCTCTCTTGGATTAGCTT	3720
DB	3455	CTAATCCAAATAAATAAAGTCAGACACANTCCACGGCTATATCTCTCTTGGATTAGCTT	3514
QY	3721	CTGCAAGTTTCATCAGCTTCTCTCCCTAATTTTACGGTTCAACAAACCTTCGTCTCAATA	3780
DB	3515	CTGCAAGTTTCATCAGCTTCTCTCCCTAATTTTACGGTTTCACAAACCTTCGTCTCAATA	3574
QY	3781	ACCGTTTGGTATAGAACCTTCTGGAGCATGCTCTTACGATCCCAACAGGTGGCTTCCA	3840
DB	3575	ACCGTTTGGTATAGAACCTTCTGGAGCATTGTCTTTCAGTCCCAACAGGTGGCTTCCA	3634
QY	3841	TGGCTCTAAGACCTTTTGATTGGCCCAAAACAGGAAGTGCCTTCCAAAGTCACAGAAACCAA	3900
DB	3635	TGGCTCTAAGACCTTTTGATTGGCCCAAAACAGGAAGTGCCTTCCAAAGTCACAGAAACCAA	3694
QY	3901	CACCTGTTTGGTCAACCAACAANTTCAAGAGCTCTCCATCACAAATCCAAATTCGATACCCA	3960
DB	3695	CACCTGTTTGGTCAACCAACAANTTCAAGAGCTCTCCATCACAAATCCAAATTCGATACCCA	3754
QY	3961	GCAACTTTTCAGTTGCTCCAGATGTAGCACTTTATACCAACCAACCGTCGACGACGAGTT	4020
DB	3755	GCACTTTTCAGTTGCTCCAGATGTAGCACTTTATACCAACCAACCGTCGACGACGAGTT	3814
QY	4021	GGTAGACTCAGATTTGTGTCTTTATAGCTCTCCGGAATAGACTTTTGGACGAGTACACCA	4080
DB	3815	GGTAGACTCAGATTTGTGTCTTTATAGCTCTCCGGAATAGACTTTTGGACGAGTACACCA	3874
QY	4081	GGCCCAACGAGTAATTAGAAGAGTCAGCCACCAAGTAGTGAATAGACCAATCGGGGGCGT	4140
DB	3875	GGCCCAACGAGTAATTAGAAGAGTCAGCCACCAAGTAGTGAATAGACCAATCGGGGGCGT	3934
QY	4141	CAGTAGTCAAGAGCCCAACAAANTTTCATGACAGGGAACCTTTTGGACATCTTCAGAAA	4200
DB	3935	CAGTAGTCAAGAGCCCAACAAANTTTCATGACAGGGAACCTTTTGGACATCTTCAGAAA	3994
QY	4201	GTTTCGTTATTCAGTAGTCAATTTGCGAGCATCAATAATGGGGAATATACCAAGACGACACAG	4260
DB	3995	GTTTCGTTATTCAGTAGTCAATTTGCGAGCATCAATAATGGGGAATATACCAAGACGACACAG	4054
QY	4261	TGGAAGTCACATCTACCAACTTTTGGCGTCTCAGAAAAAGCATAAACAGTTCTACTACCGC	4320
DB	4055	TGGAAGTCACATCTACCAACTTTTGGCGTCTCAGAAAAAGCATAAACAGTTCTACTACCGC	4114
QY	4321	CATTAGTCAAACTTTTCAATTCGCCAGTCGGAGGAAGAAAGGCACACAGCGATACTAGCAT	4380
DB	4115	CATTAGTCAAACTTTTCAATTCGCCAGTCGGAGGAAGAAAGGCACACAGCGATACTAGCAT	4174
QY	4381	TAGCGGGCAGGATGCACACTTTATCAACACAGGTCCTATAGATAACCCCTAGCGCTGGGA	4440
DB	4175	TAGCGGGCAGGATGCACACTTTATCAACACAGGTCCTATAGATAACCCCTAGCGCTGGGA	4234
QY	4441	TCAATCCTTTTGACAACTTTTCTGCCAAATCTAGGTCCTCAAAATCACATTCATTGATACCAT	4500
DB	4235	TCAATCCTTTTGACAACTTTTCTGCCAAATCTAGGTCCTCAAAATCACATTCATTGATACCAT	4294
QY	4501	TATTGTACAACTTGAGCAAGTTGTCGATCAGCTCCTCAAAATGGTCTCTGTAAACGAGTG	4560
DB	4295	TATTGTACAACTTGAGCAAGTTGTCGATCAGCTCCTCAAAATGGTCTCTGTAAACGAGTG	4354
QY	4561	ACTCAACTTGGACATTAATCTTGAGCTCAGTTCGATTGAGTGAACTTGTATCAGGTTGTGCA	4620
DB	4355	ACTCAACTTGGACATTAATCTTGAGCTCAGTTCGATTGAGTGAACTTGTATCAGGTTGTGCA	4414

QY 4621 GCTGTCAGCAGCATAGGAAACACGCGTTTCTTACCAAACTCAAGGAATTTCAAAC 4680
Db 4415 GCTGTCAGCAGCATAGGAAACACGCGTTTCTTACCAAACTCAAGGAATTTCAAAC 4474
QY 4681 CTGCAACACTTTCGCTATCCAGGTAGCAGGGAATGTCATCTTGAAGTCGGACAGTGAG 4740
Db 4475 CTGCAACACTTTCGCTATCCAGGTAGCAGGGAATGTCATCTTGAAGTCGGACAGTGAG 4534
QY 4741 TGTAGTCTTTGAGAAATCTGAAGCGGTATTTTATTCAGTGAGTCAGTCATCAGGAGA 4800
Db 4535 TGTAGTCTTTGAGAAATCTGAAGCGGTATTTTATTCAGTGAGTCAGTCATCAGGAGA 4594
QY 4801 TCCTCTAGCCGAGCGCATCTGCGCGA----- 4828
Db 4595 TCCTCTAGCCGAGCGCATCTGCGCGA----- 4654
QY 4829 ----- 4828
Db 4655 CTCGTGAAGAGGTGTTGCTGACTCATACAGGCTGNAATCGCCCATCATCCAGCCAGA 4714
QY 4829 ----- 4828
Db 4715 AAGTGAGGAGCCACGGTTGATGAGAGCTTTGTTGTAGGTGACACAGTTGGTGAATTTGA 4774
QY 4829 ----- 4828
Db 4775 ACTTTTGCTTTGCCACGAAACGCTCTGCGTTGTCGGGAAGATGCGTGATCTGATCTCTCA 4834
QY 4829 ----- 4828
Db 4835 ACTCAGCAAAAGTTCGATTTATTCAACAAGACGCGCTCCGTCAGTCAGCGTAATGCT 4894
QY 4829 ----- 4828
Db 4895 CTGCCAGTGTACAAACCAATTAACCAATCTGATTAGAAAACTCATCGAGCATCAAAATG 4954
QY 4829 ----- 4828
Db 4955 AAAGTCAATTTATTCATATCAGGATTAATCAATACATATTTTGGAAAGCCGTTCTG 5014
QY 4829 ----- 4828
Db 5015 TAATGAAGGAGAAACTCACCGAGGAGTTTCCATPAGGATGGCAAGATTCCTGGTATCGGTC 5074
QY 4829 ----- 4828
Db 5075 TGGCATCCGACTCGTCCAACTCAATACAACTTAAATTTCCCTCGTCAAAATAG 5134
QY 4829 ----- 4828
Db 5135 GTTATCAAGTGAGAAATCACCATGAGTGACGACTGAAATCCGGTGAGAAATGGCAAAAGCTT 5194
QY 4829 ----- 4828
Db 5195 ATGCAATTTCTTCAGACTTGTTCACAGCCAGCCATTAAGCTTCGTCATCAAAATCACT 5254
QY 4829 ----- 4828
Db 5255 GGCATCAACCAACCGTTATTCTATTCTGATTTGGCTGAGCGAGACGAAATACGGGATC 5314
QY 4829 ----- 4828
Db 5315 GCTGTTAAAGGACAAATTAACAAACAGGAATCGAATGCAACCGGCGAGGAACACTGCGCAG 5374
QY 4829 ----- 4828
Db 5375 CGCATCAACAAATTTTCCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTT 5434
QY 4829 ----- 4828
Db 5435 CCCGGGATCGAGTGTGATTAACCATGATCATCAGGAGTACGGATAAAATGCTTGAT 5494
QY 4829 ----- 4828

Db 5495 GGTGGAAGAGGAGCATAAATTCGTGAGCCAGTTTAGTCTGACCATCTCATCTGTAACATC 5554
QY 4829 ----- 4828
Db 5555 ATTGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGCGCATCGGCTTCCCAT 5614
QY 4829 ----- 4828
Db 5615 CAATCGATAGATTGTGCAACCTGATTGCCGACATTTATCGGAGCCCATTTATACCCATA 5674
QY 4829 ----- 4828
Db 5675 TAAATCAGATCCATGTTTGGAAATTTAATCGCGGCTCGAGCAAGACGTTTCCCGTTGAAT 5734
QY 4829 ----- 4828
Db 5735 ATGGCTCATACACCCCTTGTATTACTGTTATGTAAAGCAGACAGTTTTATTGTTCAATGA 5794
QY 4829 ----- 4828
Db 5795 TGATATATTTTATCTTGTGCAATGTAATCAGAGATTTTGAGACACAACGTGGCTTTC 5854
QY 4829 ----- 4828
Db 5855 CCCCCCCCCCTGAGTCTGCGCATCACCGCGCCACAGTCTGCTGCGCTATAT 5914
QY 4880 CGCCGACATCACCGATGCGGGAAGATCGGCTCGCACTTCGGGCTCATGAGCGCTTGT 4939
Db 5915 CGCCGACATCACCGATGCGGGAAGATCGGCTCGCACTTCGGGCTCATGAGCGCTTGT 5974
QY 4940 CGGCTGCTATGTTGCGAGCCCGCTGCGCGGAGCTGTTGGCGGCACTCTCTTGA 4999
Db 5975 CGGCTGCTATGTTGCGAGCCCGCTGCGCGGAGCTGTTGGCGGCACTCTCTTGA 6034
QY 5000 TGACCATTTCTTGGCGCGGCTGCTCAACGCGCTCAACCTACTACTGCGCTCTTCT 5059
Db 6035 TGACCATTTCTTGGCGCGGCTGCTCAACGCGCTCAACCTACTACTGCGCTCTTCT 6094
QY 5060 AATGAGGAGTGCATTAAGGAGAGCGTCAGTATCTATGATTGGAAGTATGGAATGCT 5119
Db 6095 AATGAGGAGTGCATTAAGGAGAGCGTCAGTATCTATGATTGGAAGTATGGAATGCT 6154
QY 5120 GATACCCGCAATTTCTTCTGAGTCTTCTGAGTCTTCTTATCAGATTATGCCCACTAAAGCAAC 5179
Db 6155 GATACCCGCAATTTCTTCTGAGTCTTCTGAGTCTTCTTATCAGATTATGCCCACTAAAGCAAC 6214
QY 5180 CGGAGGAGGAGATTTCTGTAATTTCTCTGACTTTTGGTTCATCAGTAGTCTGAACTG 5239
Db 6215 CGGAGGAGGAGATTTCTGTAATTTCTCTGACTTTTGGTTCATCAGTAGTCTGAACTG 6274
QY 5240 TGAGACTATCTCGGTTATGACAGCAGAAATGCTCTTCTTGAGAGCAGTAATGAGTCCC 5299
Db 6275 TGAGACTATCTCGGTTATGACAGCAGAAATGCTCTTCTTGAGAGCAGTAATGAGTCCC 6334
QY 5300 ACCAATAAGAAATCTTGTATCAGGAACAACTCTTGTTCGAACTTTTTCGGTGCC 5359
Db 6335 ACCAATAAGAAATCTTGTATCAGGAACAACTCTTGTTCGAACTTTTTCGGTGCC 6394
QY 5360 TTGAACATAAATCTAGAGTGGATATGTCGGGTAGGAATGAGAGCGGCAATCTTACC 5419
Db 6395 TTGAACATAAATCTAGAGTGGATATGTCGGGTAGGAATGAGAGCGGCAATCTTACC 6454
QY 5420 TTTCGACCTTCAAGAGATGTTAGGTTTCTGATCTGATGCAACTTCAGTGCAAC 5479
Db 6455 TTTCGACCTTCAAGAGATGTTAGGTTTCTGATCTGATGCAACTTCAGTGCAAC 6514
QY 5480 GTTGCTATTTCTCAACACCATTCGGAATCCAGAGAAATCAAGTCTGTTGCTACTATT 5539
Db 6515 GTTGCTATTTCTCAACACCATTCGGAATCCAGAGAAATCAAGTCTGTTGCTACTATT 6574
QY 5540 GATCCAGCCAGTGGCTTTTGAACCTGCAATAGTGTGCTCGTGTGTTTGAAGTCACTTT 5599

Db 6575 GATCAAGCCAGTGGGTCTTTGAAACTGACAAATAGTGTGCTGTTGTTTGGAGTCACTTT 6634
Qy 5600 TGTATGAATAAATCTAGTCTTTTGATCTAAATAATCTTTGACAGCAAGGCGGATAAATACC 5659
Db 6635 TGTATGAATAAATCTAGTCTTTTGATCTAAATAATCTTTGACAGCAAGGCGGATAAATACC 6694
Qy 5660 CAATCTAAATCTCTTTTAAACGTTTAAAGGCAAGATATGTCCTGCTGATTTAAACCCC 5719
Db 6695 CAATCTAAATCTCTTTTAAACGTTTAAAGGCAAGATATGTCCTGCTGATTTAAACCCC 6754
Qy 5720 AAATCAGCTCGTGTGATCTCTCATCACTTTGAGGGCACTATCTTTGTTTGTAGAGAAAT 5779
Db 6755 AAATCAGCTCGTGTGATCTCTCATCACTTTGAGGGCACTATCTTTGTTTGTAGAGAAAT 6814
Qy 5780 TTGGGAGATCGATATCGAGAAAAGGTAAGCTGATTTTAAACGTTGAAATTTATCTCAA 5839
Db 6815 TTGGGAGATCGATATCGAGAAAAGGTAAGCTGATTTTAAACGTTGAAATTTATCTCAA 6874
Qy 5840 GATCTCTGCTCGCGCTTTTGGTGATCAGCGTGAAACCTCTGACATACGACTCCCG 5899
Db 6875 GATCTCTGCTCGCGCTTTTGGTGATCAGCGTGAAACCTCTGACATACGACTCCCG 6934
Qy 5900 GAGACGCTCACAGCTTGTCTGTAGCGGATGCCGGAGCAGACAAGCCGCTCAGGGCGCG 5959
Db 6935 GAGACGCTCACAGCTTGTCTGTAGCGGATGCCGGAGCAGACAAGCCGCTCAGGGCGCG 6994
Qy 5960 TCAGCGGCTTGGCGGTGTGCGGGCGCAGCCATGACCCAGTCACGTACAGTACGAGGGA 6019
Db 6995 TCAGCGGCTTGGCGGTGTGCGGGCGCAGCCATGACCCAGTCACGTACAGTACGAGGGA 7054
Qy 6020 GTGTATCTGGCTTAACTATGCGGCATCAGAGCAGATTTGTACTGAGAGTGACCATATGC 6079
Db 7055 GTGTATCTGGCTTAACTATGCGGCATCAGAGCAGATTTGTACTGAGAGTGACCATATGC 7114
Qy 6080 GGTGTGAATACCGCACAGATCGTAAAGAGAAAATACCGATCAGGCGCTCTTCCGCTT 6139
Db 7115 GGTGTGAATACCGCACAGATCGTAAAGAGAAAATACCGATCAGGCGCTCTTCCGCTT 7174
Qy 6140 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGGCGGTAACAGTCACT 6199
Db 7175 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGGCGGTAACAGTCACT 7234
Qy 6200 CAAAGCGGTATACGGTTATCCACAGATCAGGGGTAACGCAAGAAAGACATGTGAG 6259
Db 7235 CAAAGCGGTATACGGTTATCCACAGATCAGGGGTAACGCAAGAAAGACATGTGAG 7294
Qy 6260 CAAAGCGGTATACGGTTATCCACAGATCAGGGGTAACGCAAGAAAGACATGTGAG 6319
Db 7295 CAAAGCGGTATACGGTTATCCACAGATCAGGGGTAACGCAAGAAAGACATGTGAG 7354
Qy 6320 GCTCGCGCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC 6379
Db 7355 GCTCGCGCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC 7414
Qy 6380 CGACAGGACTATAAGATACCAAGGCTTTCCCTCGGAAGCTCCCTCGTGGCTCTCCTG 6439
Db 7415 CGACAGGACTATAAGATACCAAGGCTTTCCCTCGGAAGCTCCCTCGTGGCTCTCCTG 7474
Qy 6440 TTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTCGGGAAGCTGGGCG 6499
Db 7475 TTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTCGGGAAGCTGGGCG 7534
Qy 6500 TTCTCATAGCTCACGCTGTAGTATCTCAGTTCGGTGTAGTGTGCTTCCCTCAAGCTGG 6559
Db 7535 TTCTCATAGCTCACGCTGTAGTATCTCAGTTCGGTGTAGTGTGCTTCCCTCAAGCTGG 7594
Qy 6560 GCTGTGTACGAAACCCCGCTTACGCGGACGCTGCGCTTATCCCGTAACTATCTGTC 6619
Db 7595 GCTGTGTACGAAACCCCGCTTACGCGGACGCTGCGCTTATCCCGTAACTATCTGTC 7654
Qy 6620 TTGAGTCCAAACCCCGTAAAGACAGACTTATCGCCACTGGCAGCAGCACTGGTAAACAGGA 6679
Db 7655 TTGAGTCCAAACCCCGTAAAGACAGACTTATCGCCACTGGCAGCAGCACTGGTAAACAGGA 7714

Qy 6680 TTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTTGAAGTGTGGCCCTAACTACG 6739
Db 7715 TTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTTGAAGTGTGGCCCTAACTACG 7774
Qy 6740 GCTACTAGAAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACTTTGGGAA 6799
Db 7775 GCTACTAGAAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACTTTGGGAA 7834
Qy 6800 AAAGATTGGTAGCTCTTGATCCGGCAAAACAAACACCGCTGCTGAGCGGTGTTTTTGG 6859
Db 7835 AAAGATTGGTAGCTCTTGATCCGGCAAAACAAACACCGCTGCTGAGCGGTGTTTTTGG 7894
Qy 6860 TTGTCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCTTTTGAATCTTTT 6919
Db 7895 TTGTCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCTTTTGAATCTTTT 7954
Qy 6920 CTACGGGCTGTGACGCTCAGTGTGAACGAAAACTCAGCTTAAAGGATTTTGGTCAAGAGAT 6979
Db 7955 CTACGGGCTGTGACGCTCAGTGTGAACGAAAACTCAGCTTAAAGGATTTTGGTCAAGAGAT 8014
Qy 6980 TATCAAAAAGGATCTTCACTAGATCTTTTAAATTTAAAAATGAAAGTTTAAATCAATCT 7039
Db 8015 TATCAAAAAGGATCTTCACTAGATCTTTTAAATTTAAAAATGAAAGTTTAAATCAATCT 8074
Qy 7040 AAAATATATAGTAAACTTTGCTCTGACAGTTTACCAATGCTTAAATCAGTAGGACCTTA 7099
Db 8075 AAAATATATAGTAAACTTTGCTCTGACAGTTTACCAATGCTTAAATCAGTAGGACCTTA 8134
Qy 7100 TCTCAGGATCTCTTATTTTGGTTTCAATAGTTCCTGACTCCCGCTGCTGATATAA 7159
Db 8135 TCTCAGGATCTCTTATTTTGGTTTCAATAGTTCCTGACTCCCGCTGCTGATATAA 8194
Qy 7160 CTACGATACGGGAGGCTTACCATCTGCGCCAGTCTGCAATGATACCGCAGAGCCAC 7219
Db 8195 CTACGATACGGGAGGCTTACCATCTGCGCCAGTCTGCAATGATACCGCAGAGCCAC 8254
Qy 7220 GCTCAGCGGCTCAGATTTATACGAAATAAACGACGCGGAGGCGGAGCGGAGAA 7279
Db 8255 GCTCAGCGGCTCAGATTTATACGAAATAAACGACGCGGAGGCGGAGCGGAGAA 8314
Qy 7280 GTGTCCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAAGCTAGAG 7339
Db 8315 GTGTCCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAAGCTAGAG 8374
Qy 7340 TAAAGTAGTTCGCCAGTTTATAGTTTGGCAACGTTGTGCAATGCTGCAAGGATCGTGG 7399
Db 8375 TAAAGTAGTTCGCCAGTTTATAGTTTGGCAACGTTGTGCAATGCTGCAAGGATCGTGG 8434
Qy 7400 TGTCAGCTCGTGTGTTGGTATGGCTTCAATTCAGCTCCGCTTCCACAGATCAAGCGAG 7459
Db 8435 TGTCAGCTCGTGTGTTGGTATGGCTTCAATTCAGCTCCGCTTCCACAGATCAAGCGAG 8494
Qy 7460 TTACATCATCCCATGTTGTGCAAAAAAGCGTTAGCTCCTTCCGCTCCTCCGATCGTTG 7519
Db 8495 TTACATCATCCCATGTTGTGCAAAAAAGCGTTAGCTCCTTCCGCTCCTCCGATCGTTG 8554
Qy 7520 TCAGAAATAAGTTGGCGCAGTGTATCATCTCATGTTATGGCAGCACTGCAATAATCTC 7579
Db 8555 TCAGAAATAAGTTGGCGCAGTGTATCATCTCATGTTATGGCAGCACTGCAATAATCTC 8614
Qy 7580 TTACTGTCAATCCCATCGTAAAGTGTCTTCTGTGACTGGTGTACTCAACCAAGTCAAT 7639
Db 8615 TTACTGTCAATCCCATCGTAAAGTGTCTTCTGTGACTGGTGTACTCAACCAAGTCAAT 8674
Qy 7640 TCTGAGAAATAGTGTATCGCGGACCGAGTTGCTCTTCCCGCGGTCAACGCGGATATA 7699
Db 8675 TCTGAGAAATAGTGTATCGCGGACCGAGTTGCTCTTCCCGCGGTCAACGCGGATATA 8734
Qy 7700 CCGCGCACATAGCAGAACTTTAAAGTGTCTATCAATGGAAGAAAGTCTTCCGGGGGAA 7759
Db 8735 CCGCGCACATAGCAGAACTTTAAAGTGTCTATCAATGGAAGAAAGTCTTCCGGGGGAA 8794

181 AGCCAGATTATGGGCTTGATGGAGCTCGTCAATTCCTTCTATTAGGCTACTA 240
241 ACACCATGACATTTATTAGCCTGTCTATCTGGCCCCCTCGCGAGGTTCAATGTTGTTA 300
241 ACACCATGACATTTATTAGCCTGTCTATCTGGCCCCCTCGCGAGGTTCAATGTTGTTA 300
301 TTTCCGATCAACAACTCCGATTAACACCGGAACATCACTCCAGATGAGGCTTTCTG 360
301 TTTCCGATCAACAACTCCGATTAACACCGGAACATCACTCCAGATGAGGCTTTCTG 360
361 AGTGGGGTCAAAATAGTTTTCATGTTCCCAAAATGGCCCCAAACTGACAGTTTAAACGCT 420
361 AGTGGGGTCAAAATAGTTTTCATGTTCCCAAAATGGCCCCAAACTGACAGTTTAAACGCT 420
421 GTCTTGGAACTTAATGACAAAAGCGTGATCTCATCCAGATCAACTTAAGTTTGGTTGG 480
421 GTCTTGGAACTTAATGACAAAAGCGTGATCTCATCCAGATCAACTTAAGTTTGGTTGG 480
481 TTGAAATCTTAACGGCCAGTTGGTCAAAAAGAAACACTTCCAAAAGTCCGCAATCCGTTTGT 540
481 TTGAAATCTTAACGGCCAGTTGGTCAAAAAGAAACACTTCCAAAAGTCCGCAATCCGTTTGT 540
541 CTGTTGGTATGATTAAGATGCAATGCTCAAAAATTAATCTCAATTAAGTTAGCGAGTCT 600
541 CTGTTGGTATGATTAAGATGCAATGCTCAAAAATTAATCTCAATTAAGTTAGCGAGTCT 600
601 CTCTATCCCTTCTGAACCCCGCTGCACTGTCGGAACGCAAAATGGGAAACACCCGCT 660
601 CTCTATCCCTTCTGAACCCCGCTGCACTGTCGGAACGCAAAATGGGAAACACCCGCT 660
661 TTTTGGATGATTAATGCAATGCTCCCAATGATGCTTCCAAAGATCTGTTGGGAATACT 720
661 TTTTGGATGATTAATGCAATGCTCCCAATGATGCTTCCAAAGATCTGTTGGGAATACT 720
721 GCTGATACCTTAAAGTTCAATGATCAAAATTAATGCTTCAACCCCTACTTGACAGCAAT 780
721 GCTGATACCTTAAAGTTCAATGATCAAAATTAATGCTTCAACCCCTACTTGACAGCAAT 780
781 ATATAACAGAGGAAGCTGCTGCTTAAACCTTTTTTTTTTATCATCAATATTAGCTT 840
781 ATATAACAGAGGAAGCTGCTGCTTAAACCTTTTTTTTTTATCATCAATATTAGCTT 840
841 ACTTTCAATATGCACTGGTTCCAAATGACAGCTTTGATTTTAACGACTTTTAACGA 900
841 ACTTTCAATATGCACTGGTTCCAAATGACAGCTTTGATTTTAACGACTTTTAACGA 900
901 CAACCTTGAGAGATCAAAAACCACTAATTTATGGAAGATCCAAACGATGAGATTTCTT 960
901 CAACCTTGAGAGATCAAAAACCACTAATTTATGGAAGATCCAAACGATGAGATTTCTT 960
961 TCAATTTTACTGCACTTTTATTCGAGCATCTCCGCAATAGCTGCTCCAGTCAACACT 1020
961 TCAATTTTACTGCACTTTTATTCGAGCATCTCCGCAATAGCTGCTCCAGTCAACACT 1020
1021 ACAACAGAGATGAAACGCAAAATTCGGCTGAAGCTGTCATCGGTTACTCAGATTTA 1080
1021 ACAACAGAGATGAAACGCAAAATTCGGCTGAAGCTGTCATCGGTTACTCAGATTTA 1080
1081 GAAGGGGATTCGATGTTCTGTTTGGCATTTTCCAAACAGCACAAATAACGGGTTATTG 1140
1081 GAAGGGGATTCGATGTTCTGTTTGGCATTTTCCAAACAGCACAAATAACGGGTTATTG 1140
1141 TTTTAAATTAATTAATTCGCAAGATTTGCTGCTTAAAGAAAGGGGATCTCTCGAGAAA 1200
1141 TTTTAAATTAATTAATTCGCAAGATTTGCTGCTTAAAGAAAGGGGATCTCTCGAGAAA 1200
1201 AGAGAGGCTGAAGCCCGGAGGATTAATGTTGGCTTTCCCGGTGTCACCGCCAGGAGTGCAG 1260
1201 AGAGAGGCTGAAGCTTTA----- 1217
1261 GCGCCCGGAGGAGGATTAATGTTGGCTTTCCCGGTGTCACCGCCAGGAGTGCAG 1320
1218 ----- 1217

1321 AGAGGTTGCTGTTTGTGATGACAGTGTCCGGGGATTCCTGGTGTCTCCACCCCATGGC 1380
1218 ----- 1217
1381 ATCGAGACACTCAAGAAGAAGATGTCCCTTTCTAACTAGTGGCCGTAGAAATTCCTAGGG 1440
1218 ----- 1234
1441 CGGCCGGAATTAATTCGGCTTAGACATGACTGTTCTCAGTTCAGTTGGGCACCTTAGC 1500
1235 CGGCCGGAATTAATTCGGCTTAGACATGACTGTTCTCAGTTCAGTTGGGCACCTTAGC 1294
1501 AGAAGACCCGCTCTGTAGATTCTAATCAAAGAGGATGTCAAGATGCCATTTCCCTGAGAG 1560
1295 AGAAGACCCGCTCTGTAGATTCTAATCAAAGAGGATGTCAAGATGCCATTTCCCTGAGAG 1354
1561 ATGAGGCTTCATTTTGTATACATTTTATTTTGAACCTATATAGTATAGGATTTTTTTT 1620
1355 ATGAGGCTTCATTTTGTATACATTTTATTTTGAACCTATATAGTATAGGATTTTTTTT 1414
1621 GTCATTTTCTTCTCTCGTAGAGCTTCTCTCATCAGCCTATCTCCAGCTGATGAA 1680
1415 GTCATTTTCTTCTCTCGTAGAGCTTCTCTCATCAGCCTATCTCCAGCTGATGAA 1474
1681 TATCTTTGTGTAAGGTTTGGGAAAATCATTCGAGTTTGAATGTTTCTTGTGATTTTCC 1740
1475 TATCTTTGTGTAAGGTTTGGGAAAATCATTCGAGTTTGAATGTTTCTTGTGATTTTCC 1534
1741 ACTCCTCTTCAGAGTACAGAGATTAAGTGAGAGTTGCTTTGTGCAAGCTTATCGATAA 1800
1535 ACTCCTCTTCAGAGTACAGAGATTAAGTGAGAGTTGCTTTGTGCAAGCTTATCGATAA 1594
1801 GCTTTAATGCGGTAGTTTATCAAGTTAAATTCGTAACAGTCAGGCAACCGTGTATGAA 1860
1595 GCTTTAATGCGGTAGTTTATCAAGTTAAATTCGTAACAGTCAGGCAACCGTGTATGAA 1654
1861 ATCTAACAATGCGCTCATCGTCACTCTCGGACCCGTCACCCCTGATGCTGAGGATAG 1920
1655 ATCTAACAATGCGCTCATCGTCACTCTCGGACCCGTCACCCCTGATGCTGAGGATAG 1714
1921 CTTGGTTATGCGGCTACTGCGGGCCCTCTTGGGGATATCGTCCATTCGACAGCATCGC 1980
1715 CTTGGTTATGCGGCTACTGCGGGCCCTCTTGGGGATATCGTCCATTCGACAGCATCGC 1774
1981 CAGTCACTATGCGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCGCT 2040
1775 CAGTCACTATGCGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCGCT 1834
2041 TCTCGGAGCACTGTCGAGCCGCTTTGCGCGCCGCGAGTCTCGCTTCTGCTACTTGG 2100
1835 TCTCGGAGCACTGTCGAGCCGCTTTGCGCGCCGCGAGTCTCGCTTCTGCTACTTGG 1894
2101 AGCCACTATCGACTACCGGATCATGCGGACCAACCCGCTCCTGTGATCTATCGAATCTA 2160
1895 AGCCACTATCGACTACCGGATCATGCGGACCAACCCGCTCCTGTGATCTATCGAATCTA 1954
2161 AATGTAAGTTAAATCTTAATAATTAAGTCCGATTTCTCCATACGAACTTAA 2220
1955 AATGTAAGTTAAATCTTAATAATTAAGTCCGATTTCTCCATACGAACTTAA 2014
2221 CAGCATTGCGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAA 2280
2015 CAGCATTGCGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAA 2074
2281 TATGTTTCACTCCCTCAGGATTAAGTCTTGTGAAGTGAAGTCTTCTGGAAGTTGCGAG 2340
2075 TATGTTTCACTCCCTCAGGATTAAGTCTTGTGAAGTGAAGTCTTCTGGAAGTTGCGAG 2134
2341 TGTTAACCTCGCTGATTTGACGGGCATATCCGTACGTTGGCAAGTGTGGTTGCTACCGG 2400
2135 TGTTAACCTCGCTGATTTGACGGGCATATCCGTACGTTGGCAAGTGTGGTTGCTACCGG 2194

4355 ACTCAACTTGCACATTAACCTTGAAGCTCAGTCGATTGAGTGAACTTGATCAGGTTGTGCA 4414 Db
4621 GCTGTCAGCAGCATAGGAAACACCGCTTTTCCTACCAAACTCAAGAAATTAACAAC 4680 Qy
4415 GCTGTCAGCAGCATAGGAAACACCGCTTTTCCTACCAAACTCAAGAAATTAACAAC 4474 Db
4681 CTGCAACACTTTCGCTATGTCAGGTAGCAAGGAAATGTCTATCTTGAAGTCGGACAGTGAG 4740 Qy
4475 CTGCAACACTTTCGCTATGTCAGGTAGCAAGGAAATGTCTATCTTGAAGTCGGACAGTGAG 4534 Db
4741 TGTAGCTTTGAGAAATTCCTGAAGCCGTATTTTATATATCAGTGAGTCAGTCATCAGGAGA 4800 Qy
4535 TGTAGCTTTGAGAAATTCCTGAAGCCGTATTTTATATATCAGTGAGTCAGTCATCAGGAGA 4594 Db
4801 TCCTCTACGCGGACGCTCGTGCGCA- 4828 Qy
4595 TCCTCTACGCGGACGCTCGTGCGCACTGTCAGGCGGGGGGGCGCTGAGGTCGTC 4654 Db
4829 - 4828 Qy
4655 CTGCTGAAGAGGTGTGTGACTCATACAGGCCCTGAATGCCCCCATCATCCAGCCAGA 4714 Db
4829 - 4828 Qy
4715 AAGTGAGGAGCCACGGTTGATGAGAGCTTTGTTGTAGGTGGAACAGTTGGTGAATTTGA 4774 Db
4829 - 4828 Qy
4775 ACTTTTGTCTTCCACGGAACGGTCTCGTGTGTCGGGAAGATGCGTGATCTGATCCTTCA 4834 Db
4829 - 4828 Qy
4835 ACTCAGAAAAGTTTCGATTTATTTCAAAAAGCCGCTCCGTCAGTCAGCGTAATGCT 4894 Db
4829 - 4828 Qy
4895 CTGCCAGTTTACACCAATTAACCAATTCGATTAGAAAACATCATCGAGCATCAATG 4954 Db
4829 - 4828 Qy
4955 AAACTGCAATTTATTCATATCAGGATTTATCAATACCATATTTTGAAAAAGCCGTTTCTG 5014 Db
4829 - 4828 Qy
5015 TAATGAAGGAGAAAACCTCACGAGGAGTTCCATAGGATGGCAAGATCCTGGTATCGTGC 5074 Db
4829 - 4828 Qy
5075 TCGATTCCGACTCGTCCCAACATCAATCAACCTATTATTTCCCTCGTCAAAAATAAG 5134 Db
4829 - 4828 Qy
5135 GTTATCAAGTGAGAAATCAACATGAGTGCAGTGAATCCGGTGAGAAATGGCAAAAGCTT 5194 Db
4829 - 4828 Qy
5195 ATGCATTTCTTCCAGACTTGTTCACAGGCCAGCCATTAACCTCGTCATCAAAATCACT 5254 Db
4829 - 4828 Qy
5255 CGCATCAACCAACCGTTATTTCATTGCTGATTCGCTGAGCGAGACGAAATACGCGATC 5314 Db
4829 - 4828 Qy
5315 GCTGTTAAAGGACAAATTAACAAACAGGAATCGAATGAACCGCGCAGGAACACTGCCAG 5374 Db
4829 - 4828 Qy
5375 CGCATCAACAATTTTACCTGAATCAGGATATTTCTTAATACTCGAATGCTGTTT 5434 Db
4829 - 4828 Qy
5435 CCGGGGATCGCAGTGGTGAGTAACCATGCAATCATCAGGATCAGGATGAAGAAATGCTTGAT 5494 Db

4829 - 4828 Qy
5495 GGTGGAGAGGCATAAATTCGTAGCCAGTTAGTCTGACCATCTCATCTGTAAATC 5554 Db
4829 - 4828 Qy
5555 ATTGGCAACGCTACCTTTGTCATGTTTTCAGAAACAACTCTGGCGCATCGGCTTCCATA 5614 Db
4829 - 4828 Qy
5615 CAATCGATAGATTGTGCACTGATTTGCCGACATTTATCGGAGCCATTTATACCCATA 5674 Db
4829 - 4828 Qy
5675 TAAATCAGCATCATGTTGGAATTTAATCGCGCCTCGAGAACGCTTTCCGTTGAAT 5734 Db
4829 - 4828 Qy
5735 ATGGCTATAAACCCCTTGTATTATGTTTATGTAAGCAGCAGTTTATTTGTTTCATGA 5794 Db
4829 - 4828 Qy
5795 TGATATATTTTATCTTGTGCAATGTAAACATCAGAGATTTTGACACACAGTGCGCTTC 5854 Db
4829 - 4828 Qy
5855 CCCCCCCCCCTCAGTCCGCAATCACCGCGCCACAGGTGCGGTTGCTGGCGCTATAT 5914 Db
4880 CGCGCATCACCGATGGGGAAGATCGGCTCCCACTTCGGGCTCATGAGCGCTTGT 4939 Qy
5915 CGCGCATCACCGATGGGGAAGATCGGCTCCCACTTCGGGCTCATGAGCGCTTGT 5974 Db
4940 CGCGCTGGTATGTTGGCAGGCCCCGGGCGGAGTGTGGGCGCCATCTCCTTGCA 4999 Qy
5975 CGCGCTGGTATGTTGGCAGGCCCCGGGCGGAGTGTGGGCGCCATCTCCTTGCA 6034 Db
5000 TGACACATTTCTTCGCGCGCGGTGCTCAACGCGCTCAACCTACTACTGCGGCTTCCT 5059 Qy
6035 TGACACATTTCTTCGCGCGCGGTGCTCAACGCGCTCAACCTACTACTGCGGCTTCCT 6094 Db
5060 AATGCAGAGTCGCATAAAGGAGAGGCTCGAGTATCTATGATTGGAAGTATGGGAATGGT 5119 Qy
6095 AATGCAGAGTCGCATAAAGGAGAGGCTCGAGTATCTATGATTGGAAGTATGGGAATGGT 6154 Db
5120 GATACCGCATTTCTCAGTGTCTTGAAGTCTCTATCAGATTTATGCCCACTAAGCAAC 5179 Qy
6155 GATACCGCATTTCTCAGTGTCTTGAAGTCTCTATCAGATTTATGCCCACTAAGCAAC 6214 Db
5180 CGGAGGAGGAGATTTTCATGTTAAATTTCTGACTTTTGGTCTCATCAGTAGACTCGAACTG 5239 Qy
6215 CGGAGGAGGAGATTTTCATGTTAAATTTCTGACTTTTGGTCTCATCAGTAGACTCGAACTG 6274 Db
5240 TGAGACTATCTCGGTTATGACAGCAAAATGTCCTTCTTGGAGACAGTAAATGAAGTCCC 5299 Qy
6275 TGAGACTATCTCGGTTATGACAGCAAAATGTCCTTCTTGGAGACAGTAAATGAAGTCCC 6334 Db
5300 ACCAATAAGAAATCCTTGTATCAGAACCAATCTTGTTCGAACTTTTTCGTTGCC 5359 Qy
6335 ACCAATAAGAAATCCTTGTATCAGAACCAATCTTGTTCGAACTTTTTCGTTGCC 6394 Db
5360 TTGAACCTATAAATGTAGAGTGGATATGTCGGTAGGAATGGAGCGGCAAAATGCTTACC 5419 Qy
6395 TTGAACCTATAAATGTAGAGTGGATATGTCGGTAGGAATGGAGCGGCAAAATGCTTACC 6454 Db
5420 TTCTGACCTTCAAGAGTATGAGGTTTGTAGTACTGATGCCCACTTCAGTGCAAC 5479 Qy
6455 TTCTGACCTTCAAGAGTATGAGGTTTGTAGTACTGATGCCCACTTCAGTGCAAC 6514 Db
5480 GTTGCTATTTTCGTTCAAAACCAATCCAGAGAAATCAAAAGTTGTTTGTCTACTATT 5539 Qy
6515 GTTGCTATTTTCGTTCAAAACCAATCCAGAGAAATCAAAAGTTGTTTGTCTACTATT 6574 Db


```
Db 8735 CCGGCCACATAGCAGAACTTTAAAGTGTCTCATCTTGGAAACCGTTCTTCGGGGCGAA 8794
Qy 7760 AACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAAACCCACTCGTGCCACCA 7819
Db 8795 AACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAAACCCACTCGTGCCACCA 8854
Qy 7820 ACTGATCTTCAGCATCTTTTACTTTCACCGCTTCTCGGTTGAGCAAAAACAGGAAGGC 7879
Db 8855 ACTGATCTTCAGCATCTTTTACTTTCACCGCTTCTCGGTTGAGCAAAAACAGGAAGGC 8914
Qy 7880 AARATGCGCAAAAGGATAGGGCGACACGGAATGTGAATCTCATCTCTTC 7939
Db 8915 AARATGCGCAAAAGGATAGGGCGACACGGAATGTGAATCTCATCTCTTC 8974
Qy 7940 TTTTTCATATTTATGAAGCACTTTATCAGGCTTATCTCATGAGCGGATACATATTG 7999
Db 8975 TTTTTCATATTTATGAAGCACTTTATCAGGCTTATCTCATGAGCGGATACATATTG 9034
Qy 8000 AATGATTTAGAAAAATAAACAATAGGGTTCGCGCACATTTCCCGAAAAGTGCCAC 8059
Db 9035 AATGATTTAGAAAAATAAACAATAGGGTTCGCGCACATTTCCCGAAAAGTGCCAC 9094
Qy 8060 CTGACGCTTAAAGAACCATTTATTCATGACATTAACCTATAAAATAGGCGTATCAGCA 8119
Db 9095 CTGACGCTTAAAGAACCATTTATTCATGACATTAACCTATAAAATAGGCGTATCAGCA 9154
Qy 8120 GGCCTTTTCGCTTCAAGAAATAATCTCATGTTTGTGACGCTTATCATGATAGCTGAC 8179
Db 9155 GGCCTTTTCGCTTCAAGAAATAATCTCATGTTTGTGACGCTTATCATGATAGCTGAC 9214
Qy 8180 TCATGTTGGTATTGTGAATAGACGAGATCGGGAACACTGAAAAATAACAGTTATTATT 8239
Db 9215 TCATGTTGGTATTGTGAATAGACGAGATCGGGAACACTGAAAAATAACAGTTATTATT 9274
Qy 8240 CG 8241
Db 9275 CG 9276

RESULT 13
ID ABZ58513 standard; DNA; 9276 BP.
XX AC ABZ58513;
XX DT 13-MAY-2003 (first entry)
XX DE Pichia pastoris plasmid pPIC9k.
XX KW Transmembrane serine protease 20; MtsP20; enzyme; endotheliasis;
XX KW cytostatic; dermatological; cardiant; vulnery; ophthalmological;
XX KW gene therapy; pPIC9k, ds.
XX OS Pichia pastoris.
XX FH Key
XX FT Location/Qualifiers
XX FT 1..948
XX FT /*tag= a
XX FT /note= "AOX1 promoter fragment"
XX FT 949..1218
XX FT /*tag= b
XX FT /note= "alpha-factor secretion signal"
XX FT 1253..1586
XX FT /*tag= c
XX FT /note= "AOX1 transcription termination region"
XX FT complement(1980..4514)
XX FT /*tag= d
XX FT /product= "HIS4"
XX FT complement(4928..5743)
XX FT /*tag= e
XX FT /note= "kanamycin resistance gene"
XX FT 6122..6879
XX FT /*tag= f
XX FT misc_feature
```

```
FT misc_feature /note= "3' AOX1 fragment"
FT complement(7288..7961)
FT /*tag= g
FT /note= "ColE1 origin"
FT 8106..8966
FT /*tag= h
FT /note= "ampicillin resistance gene"
```

W02003004681-A2.

16-JAN-2003.

03-JUL-2002; 2002WO-US021208.

03-JUL-2001; 2001US-0302939P.

(CORV-) CORVAS INT INC.

Madison EL, Ong EO;

WPI; 2003-239207/23.

New type-II membrane-type serine protease 20 polypeptides, useful for preparing a medicament for diagnosing, treating or preventing cancer, dermatological disorders, aberrant wound repairs or crest syndromes.

Example 2; Page 206-208; 216pp; English.

The present sequence is that of Pichia pastoris vector plasmid pPIC9k. The vector was used in the expression of cDNA (see ABZ58498) encoding the protease domain (see ABP72374) of novel human type II transmembrane serine protease 20 (MtsP20) in P. pastoris host cells. The invention provides MtsP20 polypeptides (see also ABP72374), the nucleic acids encoding them (see also ABZ58499), vectors, host cells, transgenic animals, probes, primers and antibodies. Also claimed are antisense oligonucleotides, double-stranded RNAs and antibodies, and their use in claimed methods of inhibiting tumour initiation, growth or progression, and of treating or preventing a disease or disorder associated with undesired and/or uncontrolled angiogenesis or neovascularisation, especially undesired angiogenesis associated with solid neoplasms, vascular malformations and cardiovascular disorders, chronic inflammatory diseases, aberrant wound repairs, circulatory disorders, crest syndromes, dermatological disorders and ocular disorders

Sequence 9276 BP; 2493 A; 2197 C; 2078 G; 2508 T; 0 U; 0 Other;

Query Match 79.6%; Score 6561.6; DB 7; Length 9276;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;

Qy 1 AGATCTAACATCCAAAGACGAAAGGTTGAATGAAACCTTTTGGCATCCGATCCACAG 60

Db 1 AGATCTAACATCCAAAGACGAAAGGTTGAATGAAACCTTTTGGCATCCGATCCACAG 60

Qy 61 GTCCATTCTCACATAGTGCACAAACGCAAGGGGGATACACTAGCAGACCGCT 120

Db 61 GTCCATTCTCACATAGTGCACAAACGCAAGGGGGATACACTAGCAGACCGCT 120

Qy 121 TGCACACGAGGACCTCCCACTCTCTCTCTCAACACACCCACTTTTGGCATCGAAAAAC 180

Db 121 TGCACACGAGGACCTCCCACTCTCTCTCTCAACACACCCACTTTTGGCATCGAAAAAC 180

Qy 181 AGCCAGTTATTGGGCTTGATTGGAGCTCGCTCATTTCCAAATTCCTTCTATTAGGCTACTA 240

Db 181 AGCCAGTTATTGGGCTTGATTGGAGCTCGCTCATTTCCAAATTCCTTCTATTAGGCTACTA 240

Qy 241 ACACCATGACTTTATTAGCTGTCTATCTGCGCCCTCGGGAGGTTTCATGTTTCTTTA 300

Db 241 ACACCATGACTTTATTAGCTGTCTATCTGCGCCCTCGGGAGGTTTCATGTTTCTTTA 300

Qy 301 TTTCCGAATGCAACAAAGCTCCGATTAACCCGAAACATCACTCCAGATCAGGGCTTCTG 360

Db 301 TTTCCGAATGCAACAAAGCTCCGATTAACCCGAAACATCACTCCAGATCAGGGCTTCTG 360

QY	361	AGTGTGGGTCAAATAGTTTCATGTTCCCAAAATGGCCCAAACTGACAGTTTAAACGGT	420	QY	1441	CGGCGCGGAATTAATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTGGGCACTTACG	1500
DB	361	AGTGTGGGTCAAATAGTTTCATGTTCCCAAAATGGCCCAAACTGACAGTTTAAACGGT	420	DB	1235	CGGCGCGGAATTAATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTGGGCACTTACG	1294
QY	421	GTCTTGAACCTAATATGACAAAAGCGTGATCTCATCCAAAGATGAACCTAAGTTGGTTGG	480	QY	1501	AGAAACCGGTCCTGCTAGATTCTTAATCAAGAGAGATGTCAGAAATGCCATTTGGCTGAGAG	1560
DB	421	GTCTTGAACCTAATATGACAAAAGCGTGATCTCATCCAAAGATGAACCTAAGTTGGTTGG	480	DB	1295	AGAAACCGGTCCTGCTAGATTCTTAATCAAGAGAGATGTCAGAAATGCCATTTGGCTGAGAG	1354
QY	481	TTGAAATGCTTAACCGGTCAGTTGGTCAAAAAGAACTTCCAAAAGTGCCTATACCGTTTGT	540	QY	1561	ATGCAGGCTTCATTTTGTATCTTTTATTTGTAACCTATATAGTATAGAGATTTT	1620
DB	481	TTGAAATGCTTAACCGGTCAGTTGGTCAAAAAGAACTTCCAAAAGTGCCTATACCGTTTGT	540	DB	1355	ATGCAGGCTTCATTTTGTATCTTTTATTTGTAACCTATATAGTATAGAGATTTT	1414
QY	541	CTTCTTTGGTATTCATTTGAGCAATGCTCAAAAATAATCTCATTAATGCTTTAGCGAGTCT	600	QY	1621	GTCAATTTGTTCTTCTGACAGAGCTTCTCCTGATCAGCTATCTCGAGCTGTATGAA	1680
DB	541	CTTCTTTGGTATTCATTTGAGCAATGCTCAAAAATAATCTCATTAATGCTTTAGCGAGTCT	600	DB	1415	GTCAATTTGTTCTTCTGACAGAGCTTCTCCTGATCAGCTATCTCGAGCTGTATGAA	1474
QY	601	CTCTATCGCTTCGAAACCGGTCACCTCTGCGGAAACGCAATGGGNAACACCGCT	660	QY	1681	TATCTTGTGGTAGGGTTTGGGAAAATCAATTCGAGTTTGTATTTTCTTGGTATTTCCC	1740
DB	601	CTCTATCGCTTCGAAACCGGTCACCTCTGCGGAAACGCAATGGGNAACACCGCT	660	DB	1475	TATCTTGTGGTAGGGTTTGGGAAAATCAATTCGAGTTTGTATTTTCTTGGTATTTCCC	1534
QY	661	TTTTGGATGATTATGCTTCCCAATTTGTATGCTTCCAAAGATCTGGTGGGAATACT	720	QY	1741	ACTCTCTTTCAGAGTACAGAAAGATTAAGTGAGAAAGTTTCGTTTGTGCAAGCTTATCGATAA	1800
DB	661	TTTTGGATGATTATGCTTCCCAATTTGTATGCTTCCAAAGATCTGGTGGGAATACT	720	DB	1535	ACTCTCTTTCAGAGTACAGAAAGATTAAGTGAGAAAGTTTCGTTTGTGCAAGCTTATCGATAA	1594
QY	721	GCTGATAGCTTAAGCTTCATGATCAAAATTTAATGTTCTTAACCCCTAATTGACAGCAAT	780	QY	1801	GCTTTAATGCGGTAGTTTATCACAGTTTAAATTTGCTAACGAGTCAGGCACCGTGTATGAA	1860
DB	721	GCTGATAGCTTAAGCTTCATGATCAAAATTTAATGTTCTTAACCCCTAATTGACAGCAAT	780	DB	1595	GCTTTAATGCGGTAGTTTATCACAGTTTAAATTTGCTAACGAGTCAGGCACCGTGTATGAA	1654
QY	781	ATATAACAGAGAGAGTCCCTGCTTAAACCTTTTATCATCATTAATTAAGCTT	840	QY	1861	ATCTAACAAATCGCTCATCGTCTCGGCAACCGTCAACCGTGGATGCTGTAGGCAATAGG	1920
DB	781	ATATAACAGAGAGAGTCCCTGCTTAAACCTTTTATCATCATTAATTAAGCTT	840	DB	1655	ATCTAACAAATCGCTCATCGTCTCGGCAACCGTCAACCGTGGATGCTGTAGGCAATAGG	1714
QY	841	ACTTTTCAATTTGCGAGTCTTCCAAATTTGACAAAGCTTTTGAATTTAACGACTTTAAACGA	900	QY	1921	CTTGGTTATGCGGTACTGCGGGCCCTTTCGCGGATATCGTCCATTCGACACATCGC	1980
DB	841	ACTTTTCAATTTGCGAGTCTTCCAAATTTGACAAAGCTTTTGAATTTAACGACTTTAAACGA	900	DB	1715	CTTGGTTATGCGGTACTGCGGGCCCTTTCGCGGATATCGTCCATTCGACACATCGC	1774
QY	901	CAACTTGAAGATCAAAACAACTAATTTATTCGAGAGTCCAAAGATGAGATTTCCCT	960	QY	1981	CAGTCACATATGCGGTGCTGTAGGCTATATGCGTTGATGCAATTTCTATGCGCACCGGT	2040
DB	901	CAACTTGAAGATCAAAACAACTAATTTATTCGAGAGTCCAAAGATGAGATTTCCCT	960	DB	1775	CAGTCACATATGCGGTGCTGTAGGCTATATGCGTTGATGCAATTTCTATGCGCACCGGT	1834
QY	961	TCAATTTTACTGCAATTTTATTCGAGCATCTTCCGATTTAGCTGCTCCAGTCAACACT	1020	QY	2041	TCTCGGACATGTCGACCGCTTTGGCCGCGCCGAGTCTCTGCTCGCTTCTGCTTACCTGG	2100
DB	961	TCAATTTTACTGCAATTTTATTCGAGCATCTTCCGATTTAGCTGCTCCAGTCAACACT	1020	DB	1835	TCTCGGACATGTCGACCGCTTTGGCCGCGCCGAGTCTCTGCTCGCTTCTGCTTACCTGG	1894
QY	1021	ACAACAGAGATGAAACGGCAAAAATTCGGCTGAAGCTGTCTATCGGTTTACTCAGATTTA	1080	QY	2101	AGCCACTATGACATACGCGATCATGGGACACACACCGCTCTGTGGATCTATCGAATCTA	2160
DB	1021	ACAACAGAGATGAAACGGCAAAAATTCGGCTGAAGCTGTCTATCGGTTTACTCAGATTTA	1080	DB	1895	AGCCACTATGACATACGCGATCATGGGACACACACCGCTCTGTGGATCTATCGAATCTA	1954
QY	1081	GAAGGGATTTTCGATTTGCTGTTTTCGCAATTTTCCAAAGCAGCAATAACGGTTTATG	1140	QY	2161	AATGTAAGTTAAATCTCTAAATTAATTAATTAAGTCCAGTTTCTCCATACGACCTTAA	2220
DB	1081	GAAGGGATTTTCGATTTGCTGTTTTCGCAATTTTCCAAAGCAGCAATAACGGTTTATG	1140	DB	1955	AATGTAAGTTAAATCTCTAAATTAATTAATTAAGTCCAGTTTCTCCATACGACCTTAA	2014
QY	1141	TTTATAAATATCTATTTGCGAGCATTTGCTGCTTAAGAAAGAGGGTATCTCTCGAGAA	1200	QY	2221	CAGCAATTCGGGTAGCATCTAGACCTTCAACAGCAGCAGATCCATCACCTGCTTGGCAA	2280
DB	1141	TTTATAAATATCTATTTGCGAGCATTTGCTGCTTAAGAAAGAGGGTATCTCTCGAGAA	1200	DB	2015	CAGCAATTCGGGTAGCATCTAGACCTTCAACAGCAGCAGATCCATCACCTGCTTGGCAA	2074
QY	1201	AGAGAGGCTGAAGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGAAACATGTATCATG	1260	QY	2281	TATGTTTTCAGTCCCTCAGGAGTTACGTTCTGTGAAGTATGAATCTTCTGGAAGTTGAG	2340
DB	1201	AGAGAGGCTGAAGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGAAACATGTATCATG	1260	DB	2075	TATGTTTTCAGTCCCTCAGGAGTTACGTTCTGTGAAGTATGAATCTTCTGGAAGTTGAG	2134
QY	1261	CCCCCGGGAGAGATAAATTTGTGGTCTCCCGGTCTCCCGCCAGCAGTGCACGGAG	1320	QY	2341	TGTTAACTCCGCTGATTTGACGGGCATATCCGTTAGTTGGCAAGTGTGGTTGTTACCGG	2400
DB	1261	CCCCCGGGAGAGATAAATTTGTGGTCTCCCGGTCTCCCGCCAGCAGTGCACGGAG	1320	DB	2135	TGTTAACTCCGCTGATTTGACGGGCATATCCGTTAGTTGGCAAGTGTGGTTGTTACCGG	2194
QY	1321	AGAGTTGCTGTTTGTATGACAGTGTCCGGGATTTCCCGTGGTCTTCCACCCCATGGCC	1380	QY	2401	AGGAGTAACTCCCAACTCTCTGAGAGTAGGCAACCAACAAACACAGATCCAGCGTGT	2460
DB	1321	AGAGTTGCTGTTTGTATGACAGTGTCCGGGATTTCCCGTGGTCTTCCACCCCATGGCC	1380	DB	2195	AGGAGTAACTCCCAACTCTCTGAGAGTAGGCAACCAACAAACACAGATCCAGCGTGT	2254
QY	1381	ATCGAAGACACTCAAGAAGAGAAATGTCCTTCTTAAGTGTGGGTAGAAATTCCTTAGGG	1440	QY	2461	GTACTTGTATCAACATAAGAAAGATTTCTCGATTTGCGAGTCAAGTGTTCAGAGCGT	2520
DB	1381	ATCGAAGACACTCAAGAAGAGAAATGTCCTTCTTAAGTGTGGGTAGAAATTCCTTAGGG	1440	DB	2255	GTACTTGTATCAACATAAGAAAGATTTCTCGATTTGCGAGTCAAGTGTTCAGAGCGT	2314
QY	1441	-----CGTAGAATTCCTTAGGG	1500	QY	2521	ACTGATGGACATTTCCAAAGCCCTGCTCGTAGTTTGCAACCGATAGGGTCTGATAGTGTG	2580

QY 4741 TGTAAGTCTTGAGAAATCTGAAGCCGATATTTTATATCAGTCAGTCAGTCATCAGGAGA 4800
Db 4535 TGTAAGTCTTGAGAAATCTGAAGCCGATATTTTATATCAGTCAGTCAGTCATCAGGAGA 4594
QY 4801 TCCTCTACGGCGGAGCATCGTGCCGA 4828
Db 4595 TCCTCTACGGCGGAGCATCGTGCCGA 4654
QY 4829 4828
Db 4655 CTCGTGAAGAAGGTGTGCTGACTCATACAGGCGTGAATCGGCCCATCATCCAGCCAGA 4714
QY 4829 4828
Db 4715 AAGTAGGGAGCCAGGTGTGATGAGAGCTTTGTGTAGGTGGACCAAGTTGGTGAATTTGA 4774
QY 4829 4828
Db 4775 ACTTTGCTTGCCACGGAACGGTCTGCGTTGTGCGGAAGATCGGTGATCTGATCCCTTCA 4834
QY 4829 4828
Db 4835 ACTCAGCAAAAGTTCGATTTATTCAACAAAGCCGCTCCGTCAGTCAGCGTAATGCT 4894
QY 4829 4828
Db 4895 CTGCCAGTGTACAAACCAATTAACCAATTTCTGATAGAAAAATCATACGAGCATCAAAATG 4954
QY 4829 4828
Db 4955 AAACTCGAATTTATTCATATCAGGATTAACAATACCATAATTTTGAAGAGCCGTTTCTG 5014
QY 4829 4828
Db 5015 TAATGAAGGAGAAAACTCACCGAGGAGTTCATPAGGATGGCAAGATTCCTGGTATCGGTC 5074
QY 4829 4828
Db 5075 TGCGATTCGGAATCGTGAACATCAATCAACACTTAAATTTCCCTCGTCATAAATAAG 5134
QY 4829 4828
Db 5135 GTTATCAAGTCAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAAATGGCAAAAGCTT 5194
QY 4829 4828
Db 5195 ATGCATTTCTTCAGACTGTGTTCAACAGGCCAGCCATACGTCGTCATCAAAATCACT 5254
QY 4829 4828
Db 5255 CGCATCAACCAACCGTTATTCATTCTGTGATGCGCTGAGCGAGACGAATAACGGGATC 5314
QY 4829 4828
Db 5315 GCTGTTAAAGGACAATTAACAAACAGGAATGGAATGCAACCGGCGGAGGAACACTGCCAG 5374
QY 4829 4828
Db 5375 CGCATCAACATATTTTCACTGAATCAGGATATTTCTTAATACCTGGAATGCTGTTTT 5434
QY 4829 4828
Db 5435 CCCGGGATCGAGTGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGAT 5494
QY 4829 4828
Db 5495 GGTGGAAGAGGCATAAATTCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTACATC 5554
QY 4829 4828
Db 5555 ATTGCAACGCTACTCTTTGCGATGTTTTCAGAAACAACTCTGGCGCATCGGGCTTCCATA 5614

QY 4829 4828
Db 5615 CAATCGATAGATTGTGCGCACCTGATTGCGCGACATATTCGCGAGCCCATTTATATACCOATA 5674
QY 4829 4828
Db 5675 TAAATCAGCATCCATGTTGGAATTTAATCGCGGCTCGAGCAAGACGTTTCCCGTTGAAT 5734
QY 4829 4828
Db 5735 ATGGCTCATACACCCCTTGATTACTGTTTATGTAAGCAGACAGTTTATTTATGTTTCATGA 5794
QY 4829 4828
Db 5795 TGATATAATTTTATCTGTGCAATGTAACATCAGAGATTTTGAGACACAACGTTGGCTTTC 5854
QY 4829 4879
Db 5855 CCCCCCCCCCTGCGAGTTCGCGATCACCGCGCCACAGGTGCGGTGCTGCGGCTATAT 5914
QY 4880 CGCGACATCAACGATGGGGAAGATCGGGCTCGCACTTCGGGCTCATGAGCGCTTCTTT 4939
Db 5915 CGCGACATCAACGATGGGGAAGATCGGGCTCGCACTTCGGGCTCATGAGCGCTTCTTT 5974
QY 4940 CGGCGTGGGTATGTTGGCAGGCCCGCTGGCGGGGAGTGTGGCGGCCCATCTCTCTTGCA 4999
Db 5975 CGGCGTGGGTATGTTGGCAGGCCCGCTGGCGGGGAGTGTGGCGGCCCATCTCTCTTGCA 6034
QY 5000 TGCAACATTCCTTGGCGGCGGTCTCAACGGCTCAACCTACTACTGCGCTGCTTCT 5059
Db 6035 TGCAACATTCCTTGGCGGCGGTCTCAACGGCTCAACCTACTACTGCGCTGCTTCT 6094
QY 5060 AATCAGAGTCGATTAAGGAGAGCGTCGAGTATCTATGATCGAAGTATGGAATGGT 5119
Db 6095 AATCAGAGTCGATTAAGGAGAGCGTCGAGTATCTATGATCGAAGTATGGAATGGT 6154
QY 5120 GATACCCCATCTTCTAGTGTCTTGAAGTCTCTATCAGATTTATGCCCAACTAAAGCAAC 5179
Db 6155 GATACCCCATCTTCTAGTGTCTTGAAGTCTCTATCAGATTTATGCCCAACTAAAGCAAC 6214
QY 5180 CGGAGGAGAGATTCATGTTAAATTTCTGACTTTTGTGTCATCAGTACTGCAACTG 5239
Db 6215 CGGAGGAGAGATTCATGTTAAATTTCTGACTTTTGTGTCATCAGTACTGCAACTG 6274
QY 5240 TGAGACTATCTCGGTATGACAGAGAAATGTCCTTCTTGAGACAGTAAATGAAGTCCC 5299
Db 6275 TGAGACTATCTCGGTATGACAGAGAAATGTCCTTCTTGAGACAGTAAATGAAGTCCC 6334
QY 5300 ACCAATAAGAAATCTTGTATCAGGAACAACTCTTGTGTTTGAACCTTTTCGGTGCC 5359
Db 6335 ACCAATAAGAAATCTTGTATCAGGAACAACTCTTGTGTTTGAACCTTTTCGGTGCC 6394
QY 5360 TTGAACATAAAATCTAGAGTGAATGTCGGGTAGGAATGAGCGGGCAAAATGCTTACC 5419
Db 6395 TTGAACATAAAATCTAGAGTGAATGTCGGGTAGGAATGAGCGGGCAAAATGCTTACC 6454
QY 5420 TTCTGGAACCTTCAAGAGTATGTAGGGTTGTAGACTGATGCGCACTTCAGTGACAAAC 5479
Db 6455 TTCTGGAACCTTCAAGAGTATGTAGGGTTGTAGACTGATGCGCACTTCAGTGACAAAC 6514
QY 5480 GTTGCTATTTCTGTTCAAAACCATTCGAAATCCAGAGAAATCAAGTGTGTTTCTACTATT 5539
Db 6515 GTTGCTATTTCTGTTCAAAACCATTCGAAATCCAGAGAAATCAAGTGTGTTTCTACTATT 6574
QY 5540 GATCCAAACCGAGTCGGTCTTGAAACTGCAATATGATGTCGTGTTTGAAGTCACTTT 5599
Db 6575 GATCCAAACCGAGTCGGTCTTGAAACTGCAATATGATGTCGTGTTTGAAGTCACTTT 6634
QY 5600 TGTATGAATAAATCTAGTCTTTGATCTTAATAATCTTGACGAGCAAGGGCGATAAATACC 5659
Db 6635 TGTATGAATAAATCTAGTCTTTGATCTTAATAATCTTGACGAGCAAGGGCGATAAATACC 6694
QY 5660 CAAATCTAAAACTCTTTTAAACCGTTTAAAGGACAAAGTATGTCGCTGTGTTTAAACCCC 5719

7775 GCTACACTAGAAAGACAGTAATTTGGTATCTGCGCTCTGCTGAGGCCAGTTACCTTCGGAA 7833

6800 AAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTTG 6859

7835 AAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTTG 7894

6860 TTGTGCAAGCAGCAGATTACCGCGCAGAAAAAAGGATCTCAAGAAGATCCCTTGCATCTTTT 6919

7895 TTGTGCAAGCAGCAGATTACCGCGCAGAAAAAAGGATCTCAAGAAGATCCCTTGCATCTTTT 7954

6920 CTACGGGTCGTGAGCTCAGTGAGAAACGAAATCTCAGTTAAGGATTTTGGTCAATGAGAT 6979

7955 CTACGGGTCGTGAGCTCAGTGAGAAACGAAATCTCAGTTAAGGATTTTGGTCAATGAGAT 8014

6980 TATCAAAAAGAGATCTTCACCTAGATCCTTTTAAATTAATAATTAAGATTTTAAATCAATCT 7039

8015 TATCAAAAAGAGATCTTCACCTAGATCCTTTTAAATTAATAATTAAGATTTTAAATCAATCT 8074

7040 AAAATGATATATGAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTA 7099

8075 AAAATGATATATGAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTA 8134

7100 TCTCAGCGATCTGTCTATTTTGGTTCATCCATAGTTGCTGACTCCCGTCTGTAGATAA 7159

8135 TCTCAGCGATCTGTCTATTTTGGTTCATCCATAGTTGCTGACTCCCGTCTGTAGATAA 8194

7160 CTACGATACGGGAGGGCTTACCATTGCGCCACAGTCTGCAATGATACCGGAGACCCAC 7219

8195 CTACGATACGGGAGGGCTTACCATTGCGCCACAGTCTGCAATGATACCGGAGACCCAC 8254

7220 GCTCACCGGCTCCAGATTTATACGAAATAAACACGACGACCGAGAGCGCAGCGCAGAA 7279

8255 GCTCACCGGCTCCAGATTTATACGAAATAAACACGACGACCGAGAGCGCAGCGCAGAA 8314

7280 GTGGCTCTGCAACTTTATCGGCTTCACCTCAGCTATTAATTTGTTCCGGGAAAGCTAGAG 7339

8315 GTGGCTCTGCAACTTTATCGGCTTCACCTCAGCTATTAATTTGTTCCGGGAAAGCTAGAG 8374

7340 TAAGTAGTTGCGCACTTAATAGTTTGGGAAAGTTGTCCTTCCATTTGCTGCAGGCATCTGG 7399

8375 TAAGTAGTTGCGCACTTAATAGTTTGGGAAAGTTGTCCTTCCATTTGCTGCAGGCATCTGG 8434

7400 TGTACGCTCGCTGTTTGGTATGCTTCAATTACGCTCCGGTTCCAAAGATCAAGGCGAG 7459

8435 TGTACGCTCGCTGTTTGGTATGCTTCAATTACGCTCCGGTTCCAAAGATCAAGGCGAG 8494

7460 TTACATGATCCCCATGTTTGCAAAAAAGCGTTAGCTCTCGGTCTCCGATCGTTG 7519

8495 TTACATGATCCCCATGTTTGCAAAAAAGCGTTAGCTCTCGGTCTCCGATCGTTG 8554

7520 TCGAAGTAAGTTGCGCGCAGTGTTATCACTCATGTTATGCGACGACTGCAATATCTC 7579

8555 TCGAAGTAAGTTGCGCGCAGTGTTATCACTCATGTTATGCGACGACTGCAATATCTC 8614

7580 TTACTGTATGCGCATCCGTAAGATGCTTTTCTGTGACTGGTAGTACTCAACCAAGTCAT 7639

8615 TTACTGTATGCGCATCCGTAAGATGCTTTTCTGTGACTGGTAGTACTCAACCAAGTCAT 8674

7640 TCTGAGAAATAGTGTATGCGGCGACCGAGTTGCTTTGCGCGGCTCACAACGGGATATA 7699

8675 TCTGAGAAATAGTGTATGCGGCGACCGAGTTGCTTTGCGCGGCTCACAACGGGATATA 8734

7700 CCGGCGCATATGACGAACTTTAAAGTGCTCATCTTGCAAAACGTTCTTCGGGGCGAA 7759

8735 CCGGCGCATATGACGAACTTTAAAGTGCTCATCTTGCAAAACGTTCTTCGGGGCGAA 8794

7760 AACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCACTCTGTCACCCA 7819

8795 AACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCACTCTGTCACCCA 8854

7820 ACTGATCTTCAGCATCTTTTATCTTCAACGAGTTTCTGGGTGAGCAAAAAACAGGAAGGC 7879

8855 ACTGATCTTCAGCATCTTTTATCTTCAACGAGTTTCTGGGTGAGCAAAAAACAGGAAGGC 8914

QY 781 ATATAACAGAGAGAGCTGCCCTGTCTTAAACCTTTTTTTTTTATCATATTATTAGCTT 840
Db 781 ATATAACAGAGAGAGCTGCCCTGTCTTAAACCTTTTTTTTTTATCATATTATTAGCTT 840
QY 841 ACTTTCATAATTGCGACTGGTTCCAAATTGCAAGCTTTTGAATTTTAAACGACTTTTAAACGA 900
Db 841 ACTTTCATAATTGCGACTGGTTCCAAATTGCAAGCTTTTGAATTTTAAACGACTTTTAAACGA 900
QY 901 CAACCTGAGAGATCAAAAACAACTAATTTTCGAGGATCCAAACGATGAGATTTCCCT 960
Db 901 CAACCTGAGAGATCAAAAACAACTAATTTTCGAGGATCCAAACGATGAGATTTCCCT 960
QY 961 TCAATTTTACTGAGTTTTTATTCGAGACTTCTCCGCAATTTAGCTGCTCCAGTCAACACT 1020
Db 961 TCAATTTTACTGAGTTTTTATTCGAGACTTCTCCGCAATTTAGCTGCTCCAGTCAACACT 1020
QY 1021 ACAACAGAGATGAACCGGCAAAATTTCCGCTGAGAGTGTCTGCTTACTCAGATTTA 1080
Db 1021 ACAACAGAGATGAACCGGCAAAATTTCCGCTGAGAGTGTCTGCTTACTCAGATTTA 1080
QY 1081 GAAGGGGATTTGCGATTTGCTGTTTGGCAATTTTCCACAGCAAAATAACCGGTTATTG 1140
Db 1081 GAAGGGGATTTGCGATTTGCTGTTTGGCAATTTTCCACAGCAAAATAACCGGTTATTG 1140
QY 1141 TTTATAAATACTACTATTGCGAGCATTTCTCTAAAGAGAGGGGTATCTCTCGAGAA 1200
Db 1141 TTTATAAATACTACTATTGCGAGCATTTCTCTAAAGAGAGGGGTATCTCTCGAGAA 1200
QY 1201 AGAGAGGCTGAAGCCCGAGCCGAGCCGAGCCGAGCCGAGGAGGAACATGATCATG 1260
Db 1201 AGAGAGGCTGAAGCTTA ----- 1217
QY 1261 GCCCCCCGGAGAGGATAAATTGTGGCTTCCCGGTGTACCGGCCAGCAGTGCACGGAG 1320
Db 1218 ----- 1217
QY 1321 AGAGGTGTGTTTTGATGACAGTGTCCGGGATTCCTGGTGTCCACCCCATGGCC 1380
Db 1218 ----- 1217
QY 1381 ATCGAGAACACTCAAGAGAGAAATGTCCCTTCTAACTAGTGGCGGTAGAAATTCCTAGGG 1440
Db 1218 ----- 1217
QY 1441 CGGCCGCGAATTAATTCGCTTAGACATGACTGTTTCTCAGTTCAAGTTGGGCACTTACG 1500
Db 1235 CGGCCGCGAATTAATTCGCTTAGACATGACTGTTTCTCAGTTCAAGTTGGGCACTTACG 1294
QY 1501 AGAGACCGGCTCTGCTAGATTTTAAATCAAGAGGATGTCAGAAATGCCATTTGCCCTGAGAG 1560
Db 1295 AGAGACCGGCTCTGCTAGATTTTAAATCAAGAGGATGTCAGAAATGCCATTTGCCCTGAGAG 1354
QY 1561 ATGCAGGCTTCATTTTGTACTTTTATTTTAACTTATATAGTATAGGATTTTTTTT 1620
Db 1355 ATGCAGGCTTCATTTTGTACTTTTATTTTAACTTATATAGTATAGGATTTTTTTT 1414
QY 1621 GTCAATTTGTTCTTCTCGTACGAGCTTCTCTGATCAGCTATCTCGCAGCTCATGAA 1680
Db 1415 GTCAATTTGTTCTTCTCGTACGAGCTTCTCTGATCAGCTATCTCGCAGCTCATGAA 1474
QY 1681 TATCTTTGTGTAGGGTTTTGGGAAATCAATTCAGTTTGAATTTTTCTTGTATTTCCT 1740
Db 1475 TATCTTTGTGTAGGGTTTTGGGAAATCAATTCAGTTTGAATTTTTCTTGTATTTCCT 1534
QY 1741 ACTCCTCTTCAGAGTACAGAGATTAAGTGAGAGTTTGTGTGCAAGCTTATCGATTA 1800
Db 1535 ACTCCTCTTCAGAGTACAGAGATTAAGTGAGAGTTTGTGTGCAAGCTTATCGATTA 1594
QY 1801 GCTTTAATGCGGTAGTTTATCAGTTTAAATTCCTAACTCAGGTCAGGACCGGTGATGAA 1860
Db 1595 GCTTTAATGCGGTAGTTTATCAGTTTAAATTCCTAACTCAGGTCAGGACCGGTGATGAA 1654

QY 1861 ATCTAAACATGGCTCATCTGTCATCTCTGGGACCGGTCAACCTGGATGCTGTAGGCATAGG 1920
Db 1655 ATCTAAACATGGCTCATCTGTCATCTCTGGGACCGGTCAACCTGGATGCTGTAGGCATAGG 1714
QY 1921 CTTTGGTTATGGCGGTACTGCGGGCTCTTTCGGGATATCGTCCATTCGACAGCATGCG 1980
Db 1715 CTTTGGTTATGGCGGTACTGCGGGCTCTTTCGGGATATCGTCCATTCGACAGCATGCG 1774
QY 1981 CAGTCACTATGGCTGCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGGGACCGGT 2040
Db 1775 CAGTCACTATGGCTGCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGGGACCGGT 1834
QY 2041 TCTCGGAGCACTGTCGACCGCTTTGGCGCGCCGAGTCTGCTGCTGCTTCTGCTACTTGG 2100
Db 1835 TCTCGGAGCACTGTCGACCGCTTTGGCGCGCCGAGTCTGCTGCTGCTTCTGCTACTTGG 1894
QY 2101 AGCCACTATGCACTAGCGCATCATGCGGACCAACACCGCTCTGCTGCTGATATGCAATCTA 2160
Db 1895 AGCCACTATGCACTAGCGCATCATGCGGACCAACACCGCTCTGCTGCTGATATGCAATCTA 1954
QY 2161 AATGTAAGTTAAATCTCTAAATTAATTAATTAAGTCCGAGTTTCTCCATACCAACCTTAA 2220
Db 1955 AATGTAAGTTAAATCTCTAAATTAATTAATTAAGTCCGAGTTTCTCCATACCAACCTTAA 2014
QY 2221 CAGCATTGCGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCCATCTGCTTGGCCAA 2280
Db 2015 CAGCATTGCGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCCATCTGCTTGGCCAA 2074
QY 2281 TATGTTTCACTCCCTCAGGAGTTACCTTTGTGAAGTATGAACCTTCTGGAAGTTGCGAG 2340
Db 2075 TATGTTTCACTCCCTCAGGAGTTACCTTTGTGAAGTATGAACCTTCTGGAAGTTGCGAG 2134
QY 2341 TGTTAACCTCGCTGATTTGACGGGCAATATCCGTACCTTGGCAAGTGTGGTTGGTACCGG 2400
Db 2135 TGTTAACCTCGCTGATTTGACGGGCAATATCCGTACCTTGGCAAGTGTGGTTGGTACCGG 2194
QY 2401 AGGAGTAATCTCCACAACTCTCTGGAGAGTAGGCACCAACAAACACAGATTCAGCGTGT 2460
Db 2195 AGGAGTAATCTCCACAACTCTCTGGAGAGTAGGCACCAACAAACACAGATTCAGCGTGT 2254
QY 2461 GTACTTGAATCAATAGAGAGAGCATTTCTCGATTTGCGAGTCAAGTGTTCAGGAGCGT 2520
Db 2255 GTACTTGAATCAATAGAGAGAGCATTTCTCGATTTGCGAGTCAAGTGTTCAGGAGCGT 2314
QY 2521 ACTGATTTGACATTTCCAAAGCCTGCTGCTGAGTGTGCAACCGATAGGTTGTAGAGTGTG 2580
Db 2315 ACTGATTTGACATTTCCAAAGCCTGCTGCTGAGTGTGCAACCGATAGGTTGTAGAGTGTG 2374
QY 2581 CAATACACTTGGCTACAATTTCAACCCCTTGGCAACTGSCACAGCTTGGTTGTGAACAGCAT 2640
Db 2375 CAATACACTTGGCTACAATTTCAACCCCTTGGCAACTGSCACAGCTTGGTTGTGAACAGCAT 2434
QY 2641 CTTCAATTTCTGGCAAGCTCCTTGTCTGTCATATCGACAGCAACAGAAATCACCTGGGAAT 2700
Db 2435 CTTCAATTTCTGGCAAGCTCCTTGTCTGTCATATCGACAGCAACAGAAATCACCTGGGAAT 2494
QY 2701 CAATACCATGTTTCAAGCTTGAGACAGAGGTTCTGAGGCAACGAAATCTGGATCAGCGTATT 2760
Db 2495 CAATACCATGTTTCAAGCTTGAGACAGAGGTTCTGAGGCAACGAAATCTGGATCAGCGTATT 2554
QY 2761 TATCAGCAATTAATAGAACTTTCAGAGGCGCCAGGCAATGTCAATACTACACAGGCGTG 2820
Db 2555 TATCAGCAATTAATAGAACTTTCAGAGGCGCCAGGCAATGTCAATACTACACAGGCGTG 2614
QY 2821 ATGTGTCAATTTGAACCATCATCTTTGGCAGCAGTAAACGAACTGGTTTCTGGACCAATA 2880
Db 2615 ATGTGTCAATTTGAACCATCATCTTTGGCAGCAGTAAACGAACTGGTTTCTGGACCAATA 2674
QY 2881 TTTTGTCACTTATAGGAACAGTTTCTGTTCCGTTAAGCCATAGCAGCTACTGCTGGGCGC 2940
Db 2675 TTTTGTCACTTATAGGAACAGTTTCTGTTCCGTTAAGCCATAGCAGCTACTGCTGGGCGC 2734
QY 2941 CTCCTGTAGCACGATACACTTTAGCACCAACCTTTGGGCAACGATAGTACTTCTGGGG 3000

Db 2735 |||||CTCCTGCTAGCAGATACACTTAGCACCAACCTTTGTGGGCAACGTAGATGACTTCTCGGG 2794
Qy 3001 |||||TAAGGGTACCACTCCTTTCTAGGTGGAGATGCAAAAACAAATTTCTTTGCAACCCAGCAACTT 3060
Db 2795 |||||TAAGGGTACCACTCCTTTCTAGGTGGAGATGCAAAAACAAATTTCTTTGCAACCCAGCAACTT 2854
Qy 3061 |||||TGGCAGGAACCCAGCATCAGGGAAGTGGGAAGCAGAAATTTGCGGTTCACCCAGGAATAT 3120
Db 2855 |||||TGGCAGGAACCCAGCATCAGGGAAGTGGGAAGCAGAAATTTGCGGTTCACCCAGGAATAT 2914
Qy 3121 |||||AGAGGCCAACTTTCTCAATAGGTCTTGCAAAACGAGAGCAGACTACACCCAGGCGCAAGTCT 3180
Db 2915 |||||AGAGGCCAACTTTCTCAATAGGTCTTGCAAAACGAGAGCAGACTACACCCAGGCGCAAGTCT 2974
Qy 3181 |||||CAACTTGCACGTCTCCGTAGTTAGCTTCAATGGAATTTCTCGACGTATCTATAGAGA 3240
Db 2975 |||||CAACTTGCACGTCTCCGTAGTTAGCTTCAATGGAATTTCTCGACGTATCTATAGAGA 3034
Qy 3241 |||||GATCAATGGCTCTCTTAACGTTATCTGGCAATTCGCAATTCGCAATTCGCTTGGGAAGGAGCTT 3300
Db 3035 |||||GATCAATGGCTCTCTTAACGTTATCTGGCAATTCGCAATTCGCAATTCGCTTGGGAAGGAGCTT 3094
Qy 3301 |||||CTAACACAGGTGTCTTCAAAAGCGACTCCATCAAACTTTGGCAGTTAGTTCTAAAGGGCTT 3360
Db 3095 |||||CTAACACAGGTGTCTTCAAAAGCGACTCCATCAAACTTTGGCAGTTAGTTCTAAAGGGCTT 3154
Qy 3361 |||||TGTACCAATTTGAGGAACATGTCGACAAATGTTGTCATTAATTCATTAATTCGTTCCG 3420
Db 3155 |||||TGTACCAATTTGAGGAACATGTCGACAAATGTTGTCATTAATTCATTAATTCGTTCCG 3214
Qy 3421 |||||TTTTCTGATAGGAGCAGCAAGGGCCTCTTCAATTTCTTTGAGGAGGCGCTTAGAAACGT 3480
Db 3215 |||||TTTTCTGATAGGAGCAGCAAGGGCCTCTTCAATTTCTTTGAGGAGGCGCTTAGAAACGT 3274
Qy 3481 |||||CAAATTTGCACAATTCATAGCACTTCAGGAAGGCACTTTCTTAGGTTTGGATTCCTT 3540
Db 3275 |||||CAAATTTGCACAATTCATAGCACTTCAGGAAGGCACTTTCTTAGGTTTGGATTCCTT 3334
Qy 3541 |||||TAGGTTTCTCTGGTATCCTGGCTTGGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3600
Db 3335 |||||TAGGTTTCTCTGGTATCCTGGCTTGGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3394
Qy 3601 |||||ACTTCATATCCAGGTTTCTCTCCACCTCTCCACCTGTCACACGTCACACGTCATCTAA 3660
Db 3395 |||||ACTTCATATCCAGGTTTCTCTCCACCTCTCCACCTGTCACACGTCACACGTCATCTAA 3454
Qy 3661 |||||CTAATGCAAAATAAATAGTCAGCAATTTCCAGGCTATATCTCTCTCTCTCTCTCTCTCTCT 3720
Db 3455 |||||CTAATGCAAAATAAATAGTCAGCAATTTCCAGGCTATATCTCTCTCTCTCTCTCTCTCTCT 3514
Qy 3721 |||||CTGCAAGTTTCACTAGCTTCT 3780
Db 3515 |||||CTGCAAGTTTCACTAGCTTCT 3574
Qy 3781 |||||ACCGTTTGGTATAGAACCTTCTGGAGCATGCTCTTAAGCATCCCAAGGTCGCTTCCA 3840
Db 3575 |||||ACCGTTTGGTATAGAACCTTCTGGAGCATGCTCTTAAGCATCCCAAGGTCGCTTCCA 3634
Qy 3841 |||||TGGCTCTAAGACCTTTGATTGGGCAAAAACAGGAAGTGGCTTCCAGTGACAGAAACCAA 3900
Db 3635 |||||TGGCTCTAAGACCTTTGATTGGGCAAAAACAGGAAGTGGCTTCCAGTGACAGAAACCAA 3694
Qy 3901 |||||CACCTGTTTGTTCACCAAAAATTTCAAGCAGTCTCCATCACAATCAAAATTCGATACCCA 3960
Db 3695 |||||CACCTGTTTGTTCACCAAAAATTTCAAGCAGTCTCCATCACAATCAAAATTCGATACCCA 3754
Qy 3961 |||||GCAACTTTTGTGTTGCTCCAGATGACCTTTATACCAAAACCGTCCAGCAGAGATT 4020
Db 3755 |||||GCAACTTTTGTGTTGCTCCAGATGACCTTTATACCAAAACCGTCCAGCAGAGATT 3814
Qy 4021 |||||GGTAGACTCCAGTTTGTCTTATAGCTCTCCGGAATAGACTTTTGGAGGAGTACACCA 4080

Db 3815 |||||GGTAGACTCCAGTTTGTGTCTTATAGCCTCCGGAATAGACTTTTGTGACGAGTACACCA 3874
Qy 4081 |||||GGCCCAACGAGTAATTAGAAGAGTACAGCCCAAAAGTAGTGAATAGACCATTCGGGGCGGT 4140
Db 3875 |||||GGCCCAACGAGTAATTAGAAGAGTACAGCCCAAAAGTAGTGAATAGACCATTCGGGGCGGT 3934
Qy 4141 |||||CAGTAGTCAAGACGCCAACAAAATTTCACTGACAGGGAACTTTTGTGACATCTTCAGAAA 4200
Db 3935 |||||CAGTAGTCAAGACGCCAACAAAATTTCACTGACAGGGAACTTTTGTGACATCTTCAGAAA 3994
Qy 4201 |||||GTTCTGTTATTCAGTACTAGTCAATTTGCCGAGCATCAATTAATGGGATTATACAGAGCAACAG 4260
Db 3995 |||||GTTCTGTTATTCAGTACTAGTCAATTTGCCGAGCATCAATTAATGGGATTATACAGAGCAACAG 4054
Qy 4261 |||||TGGAACTCACATCTTACCAACTTTTGGGTCTCAGAAAAAGCATAAACAGTTCTTACTACCGC 4320
Db 4055 |||||TGGAACTCACATCTTACCAACTTTTGGGTCTCAGAAAAAGCATAAACAGTTCTTACTACCGC 4114
Qy 4321 |||||CATTTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAGGCAACGCGATACTAGCAT 4380
Db 4115 |||||CATTTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAGGCAACGCGATACTAGCAT 4174
Qy 4381 |||||TAGCGGGCAAGGATCAAACTTTTATCAACAGGGTCTCTATAGATAACCTTAGCGCTGGGA 4440
Db 4175 |||||TAGCGGGCAAGGATCAAACTTTTATCAACAGGGTCTCTATAGATAACCTTAGCGCTGGGA 4234
Qy 4441 |||||TCATCTTTTGGCAACTCTTTCTGCAAACTAGGTCCAAATCACTTCACTTATGATACCAT 4500
Db 4235 |||||TCATCTTTTGGCAACTCTTTCTGCAAACTAGGTCCAAATCACTTCACTTATGATACCAT 4294
Qy 4501 |||||TATTTGTAACAATTTGAGCAAGTTTGTGATCAGTCTCTCAAAATTTGCTCTCTGTAACGGATG 4560
Db 4295 |||||TATTTGTAACAATTTGAGCAAGTTTGTGATCAGTCTCTCAAAATTTGCTCTCTGTAACGGATG 4354
Qy 4561 |||||ACTCAACTTGCACATTAACCTTTGAAAGTCTCAGTCTGATGAACTTGAATCAAGTTTGTGCA 4620
Db 4355 |||||ACTCAACTTGCACATTAACCTTTGAAAGTCTCAGTCTGATGAACTTGAATCAAGTTTGTGCA 4414
Qy 4621 |||||GCTGTGTGACGAGCATAGGAAACACGGCTTTTCTTACCAAACTCAAGAAATTAACAACCT 4680
Db 4415 |||||GCTGTGTGACGAGCATAGGAAACACGGCTTTTCTTACCAAACTCAAGAAATTAACAACCT 4474
Qy 4681 |||||CTGCAACACTTTCGCTATCAGGTAGCAAGGAAATGTCATCTTGAAGTCCGACAGTGAG 4740
Db 4475 |||||CTGCAACACTTTCGCTATCAGGTAGCAAGGAAATGTCATCTTGAAGTCCGACAGTGAG 4534
Qy 4741 |||||TGTAGTCTTGAAGAAATCTGGAAGCGGTATTTTATTTATCAGTGAGTCACTCATCAGGAGA 4800
Db 4535 |||||TGTAGTCTTGAAGAAATCTGGAAGCGGTATTTTATTTATCAGTGAGTCACTCATCAGGAGA 4594
Qy 4801 |||||TCCTCTACGCGGACGCTCTGCGCGA----- 4828
Db 4595 |||||TCCTCTACGCGGACGCTCTGCGCGA----- 4654
Qy 4829 |||||----- 4828
Db 4655 |||||CTCTGGAAGAGGTGTTGCTGACTATACAGGCTGAAATCGCCCATCATCCAGCAGGA 4714
Qy 4829 |||||----- 4828
Db 4715 |||||AAGTGAGGAGCCACGGTTGATGAGAGCTTTGTTGTTAGTGGACCGCTTGGTATTGGA 4774
Qy 4829 |||||----- 4828
Db 4775 |||||ACTTTTGTCTTCCACCGAACGGTCTGCGTTGTCGGGAAGATGCGTGATCTGATCTCTCA 4834
Qy 4829 |||||----- 4828
Db 4835 |||||ACTCAGCAAAAGTTCGATTATTTCAACAAAGCCGCGCTCCGTCAGTCAAGTCAAGCTAATGCT 4894
Qy 4829 |||||----- 4828
Db 4895 |||||CTGCGAGTGTATCAACCAAAATTAACCAATTTCTGATTAGAAAAAATCTCATTCGAGCATCAATG 4954

QY	4829	-----	4828	5000	TGACCAATTCCTTGGCGGGCGTCTCAACGGCCTCAACCTACTACTGGCTGCTTCT	5059
Db	4955	AAACTGCAATTTATTCATATCAGGATTCATATACCATAATTTTGAAGAGCGGTTCTG	5014	6035	TGCACCAATTCCTTGGCGGGCGTCTCAACGGCCTCAACCTACTACTGGCTGCTTCT	6094
QY	4829	-----	4828	5060	AATGACAGAGTCGCATTAAGGAGAGCGTCGATATCTATGATTGGAAGTATCGGAATGT	5119
Db	5015	TAATGAGGAGAAACTCACCAGGAGCTTCCATAGGATGGCAAGATCCTGGTATCGGTC	5074	6095	AATGACAGAGTCGCATTAAGGAGAGCGTCGATATCTATGATTGGAAGTATCGGAATGT	6154
QY	4829	-----	4828	5120	GATACCCGCAATTCCTCAGTCTCTGAGTCTCTCTATCAGATTATGCCCACTAAAGCAAC	5179
Db	5075	TGCGATTCGAGCTGTCMAACATCAATACACCTATTAAATTCCTCCCTCGTCAAAAATAG	5134	6155	GATACCCGCAATTCCTCAGTCTCTCTGAGTCTCTCTATCAGATTATGCCCACTAAAGCAAC	6214
QY	4829	-----	4828	5180	CGGAGGAGAGATTCATCGTAAATTTCTGACTTTTGGTTCATCAGTAGACTCGAACTG	5239
Db	5135	GTATCAAGTGAAGAAATCACCATGAGTCAGACTGAATCCGGTGAGAAATGGCAAAAGCTT	5194	6215	CGGAGGAGAGATTCATCGTAAATTTCTGACTTTTGGTTCATCAGTAGACTCGAACTG	6274
QY	4829	-----	4828	5240	TGAGACTATCTCGGTTATCAGCAGAAATGCTCTTCTTGGAGACAGTAAATGAAGTCC	5299
Db	5195	ATGCATTTCTTCCAGACTTGTTCAACAGGCGCAGCTTACGCTCGTCAAAAATCACT	5254	6275	TGAGACTATCTCGGTTATCAGCAGAAATGCTCTTCTTGGAGACAGTAAATGAAGTCC	6334
QY	4829	-----	4828	5300	ACCAATAAGAAATCCTTGTATCAGAAACAACTCTTGTTCGAACTTTTTCGGTGCC	5359
Db	5255	CGCATCAACCAACCGTTATTCATTCGTGATTGGCCTGAGCGAGACGAAATACGCGATC	5314	6335	ACCAATAAGAAATCCTTGTATCAGAAACAACTCTTGTTCGAACTTTTTCGGTGCC	6394
QY	4829	-----	4828	5360	TTGAACTATAAAATGTAGAGTGATGTCCGGTAGGAATGGAGCGGGCAAAATGCTTACC	5419
Db	5315	GCTGTTAAAGGACAATTAACAACAGGAATCGAATGCAACCGCGCGAGAACACTGGCAG	5374	6395	TTGAACTATAAAATGTAGAGTGATGTCCGGTAGGAATGGAGCGGGCAAAATGCTTACC	6454
QY	4829	-----	4828	5420	TTCTGACCTTCAAGAGTATGTAGAGTTGTAGATACCTGATGCCAACTTCAGTGACAAC	5479
Db	5375	CGCATCAACAATATTTTCACTGATCAGGATATCTTCTAATACCTGGATGCTGTTT	5434	6455	TTCTGACCTTCAAGAGTATGTAGAGTTGTAGATACCTGATGCCAACTTCAGTGACAAC	6514
QY	4829	-----	4828	5480	GTTCATTTTCGTTCAAACTCCGAATCCAGAAATCCAGAAATCAAAGTTGTTGCTACTATT	5539
Db	5435	CCCGGGATCGAGTGGTGAATACCATGCATCATCAGGATACGATATAAATGCTTGAT	5494	6515	GTTCATTTTCGTTCAAACTCCGAATCCAGAAATCCAGAAATCAAAGTTGTTGCTACTATT	6574
QY	4829	-----	4828	5540	GATCCAAAGCAGTCGGTCTTGAAACTGACAAATAGTGTCTGTCTGTGTGTGAGTCTATT	5599
Db	5495	GGTCGGAAGAGCATAAATTCGTCAGCCAGTTTGTAGTCTGACCATCTCATCTGTAAACATC	5554	6575	GATCCAAAGCAGTCGGTCTTGAAACTGACAAATAGTGTCTGTCTGTGTGTGAGTCTATT	6634
QY	4829	-----	4828	5600	TGTATGAATAAATCTAGTCTTTGATCTTAATAATCTTGACGAGCAAGGCGATAAATACC	5659
Db	5555	ATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGCTTCCCAT	5614	6635	TGTATGAATAAATCTAGTCTTTGATCTTAATAATCTTGACGAGCAAGGCGATAAATACC	6694
QY	4829	-----	4828	5660	CAATCTAAACCTCTTTTAAACCTTAAAGGACAAAGTATGTCTGCTGTATTAAACCC	5719
Db	5615	CAATCGATAGATTGTGCGACCTGATTCGCCGACATTAATCGAGGCCATTAATACCCATA	5674	6695	CAATCTAAACCTCTTTTAAACCTTAAAGGACAAAGTATGTCTGCTGTATTAAACCC	6754
QY	4829	-----	4828	5720	AAATCAGCTCGTAGTCTGATCCTCATCAACTTGAGGGGCACTATCTGTTTTAGAGAAAT	5779
Db	5675	TAAATCAGCATCCATGTTGGAATTTAATCGCGCCTCGAGCAAGCTTCCCGTTGAAT	5734	6755	AAATCAGCTCGTAGTCTGATCCTCATCAACTTGAGGGGCACTATCTGTTTTAGAGAAAT	6814
QY	4829	-----	4828	5780	TTGCGGAGATCGGATTCGAGAAAAAGGTACGCTGATTTTAAACGTGAAATTTATCTCAA	5839
Db	5735	ATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTATTGTTTATGA	5794	6815	TTGCGGAGATCGGATTCGAGAAAAAGGTACGCTGATTTTAAACGTGAAATTTATCTCAA	6874
QY	4829	-----	4828	5840	GATCTCTGCTCGCGCTTTTCGGTGATGACGGTGAACCTCTGACACATCGAGCTCCCG	5899
Db	5795	TGATATATTTTATCTGTGCAATGTAAATCATCAGAGTTTGTAGACACAACTGTCGAG	5854	6875	GATCTCTGCTCGCGCTTTTCGGTGATGACGGTGAACCTCTGACACATCGAGCTCCCG	6934
QY	4829	-----	4828	5900	GAGACGCTCAGACTTCTCTGTAAGCGGATGCGGGAGCAGACAGCCCGCTCAGGGCGCG	5959
Db	5855	CCCCCCCCCTCGAGTCGGCATCACCGCGCCACAGGTGCGGTTGCTGGCGCCTATAT	5914	6935	GAGACGCTCAGACTTCTCTGTAAGCGGATGCGGGAGCAGACAGCCCGCTCAGGGCGCG	6994
QY	4880	CGCGGACATCACCGATGGGAGATCGGCTCGCACTTCGGGCTCATGAGCGCTGTTT	4939	5960	TCAGCGGCTGTGCGGCTGTCCGGCGCAGCCATGACCAGTCACTAGCGATAGCGGA	6019
Db	5915	CGCCGACATCACCGATGGGAGATCGGCTCGCACTTCGGGCTCATGAGCGCTGTTT	5974	6995	TCAGCGGCTGTGCGGCTGTCCGGCGCAGCCATGACCAGTCACTAGCGATAGCGGA	7054
QY	4940	CGCGTGGGTATGCTGGCAGGCCCCGTGGCCGGGGACTGTTGGGGCCCATCTCCCTTGA	4999	6020	GTGTATCTGCTTAACTATGCGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGC	6079
Db	5975	CGCGTGGGTATGCTGGCAGGCCCCGTGGCCGGGGACTGTTGGGGCCCATCTCCCTTGA	6034	7055	GTGTATCTGCTTAACTATGCGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGC	7114
QY	5975	CGCGTGGGTATGCTGGCAGGCCCCGTGGCCGGGGACTGTTGGGGCCCATCTCCCTTGA	6034	6080	GCTGTGAAATACCGCAGACAGATGCGTAAAGGAGAAAAATACCGCATCAGGCGCTCTTCCGCTT	6139

RESULT 15
AAD09083
ID AAD09083 standard; DNA; 9776 BP.
XX AC
XX AAD09083;
DT 04-SEP-2001 (first entry)
XX
DE pUB55 expression plasmid for Sonic hedgehog protein in Pichia pastoris.
XX
KW Human; hedgehog protein; nontropic; neuroprotective; anticonvulsant;
KW cytosstatic; therapy; Alzheimer's disease; Parkinson's disease; injury;
KW Huntington's chorea; amyotrophic lateral sclerosis; multiple sclerosis;
KW nervous system aging; neurodegenerative disease; immunological disease;
KW malignant glioma; medulloblastoma; neuroectodermal tumour; cancer;
KW extracellular signalling protein; plasmid PUB55; ds.
XX
XX Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
XX WO200134654-A1.
PN
XX
PD 17-MAY-2001.
PD
XX
XX 02-NOV-2000; 200WO-USO30405.
PF
XX
XX 05-NOV-1999; 99US-0164025P.
PR
XX
XX (BIOJ) BIOGEN INC.
PA
XX
XX Strauch X;
PI
XX
XX WPI; 2001-329075/34.
DR
XX
XX Novel isolated hedgehog fusion polypeptide useful for treating
PT neurological conditions such as Alzheimer's disease, Parkinson's disease,
PT Huntington's chorea, amyotrophic lateral sclerosis, and multiple
PT sclerosis.
XX
XX Example 1; Page 62-67; 178pp; English.
PS
XX
XX The present invention relates to hedgehog fusion proteins. Hedgehog
CC proteins are a family of extracellular signalling proteins that regulate
CC various aspects of embryonic development both in vertebrates and in
CC invertebrates. Hedgehog fusion protein is useful for the prophylaxis or
CC treatment of any condition or disease state for which a hedgehog or
CC patched protein constituent is efficacious and in the diagnosis of
CC constituents or conditions of disease states in biological system or
CC specimens and for diagnostic purposes in non-physiological systems.
CC Hedgehog fusion protein is useful for treating neurological conditions
CC due to injury, aging of nervous system, including Alzheimer's disease,
CC chronic neurodegenerative diseases of the nervous system, including
CC Parkinson's disease, Huntington's chorea, amyotrophic lateral sclerosis
CC and chronic immunological diseases of nervous system including multiple
CC sclerosis and malignant gliomas, medulloblastomas, neuroectodermal
CC tumors and to specifically target medical therapies against cancers and
CC tumors which express the receptor for the protein. The present sequence
CC is pUB55 expression plasmid for sonic hedgehog protein in Pichia
CC pastoris. The plasmid contains the DNA encoding N-terminal domain of
CC human sonic hedgehog with the alpha factor prepro region as the secretion
CC signal
XX
XX Sequence 9776 BP; 2608 A; 2305 C; 2263 G; 2600 T; 0 U; 0 Other;
SQ

Query Match 77.5%; Score 6393.8; DB 4; Length 9776;
Best Local Similarity 82.9%; Pred. No. 0;
Matches 8101; Conservative 0; Mismatches 137; Indels 1535; Gaps 10

2 GATCTTAACATCAAGACCAAGGTTGTAATGAACACTTTTTCCTCCGATCCGATCCACAGG 61

QY 1142 TTATAAATCTACTATTGCGAGCATTCCTGCTAAAGAAAGAGGGTATCTCTCGAGAAA 1201
DB 1141 TTTATAAATCTACTATTGCGAGCATTCCTGCTAAAGAAAGAGGGTATCTCTCGAGAAA 1200
QY 1202 GAGAGGCTGAAGCCGAGG----- 1219
DB 1201 GATCGGACCGGAGGAGGGGTTTCGGAAGAGAGGAGCCACCCAAAAGCTGACCCCTTTAG 1260
QY 1220 -----CCGAGCCGAGGCCCCAGGCCAG 1242
DB 1261 CCTACAAGAGTTTATCCCAATGTGCGGAGAGACCCCTAGGCGCAGCGAGGTATG 1320
QY 1243 CAAGAAACAT----- 1252
DB 1321 AAGGAGATCTCCAGAACTCCGAGCGATTTAAGGAACCTCACCCCAATTACAAACCCG 1380
QY 1253 -----GTAATCATGCGCCCCCGGAGAGGATTAATTTGTGCTTCCCGGTGTC 1299
DB 1381 ACATCATATTTAAGGATGAAGAAACACCGGAGCGGACGCTGATGACTCAGAGGTGA 1440
QY 1300 AC-----CGCCAGCAGTGCACGGAGAGAGGTGCTGTTT 1334
DB 1441 AGGCAAGTTGAACGCTTTGGCCATCTCGGTGATGAACCACTGAGGAGTGAAGTGC 1500
QY 1335 TGATGACAGTGTCCGGGAGTTCCCGTGG----- 1362
DB 1501 GGGTGACCGAGGCTGGGACGAAGATGGCCACCACCTCAGAGGAGTCTCTGCACTACGAGG 1560
QY 1363 ----- 1362
DB 1561 GCCGGCAGTGACATCACACAGTCTGACCGGACCGCAGCAAGTACGGCATGCTGCGCC 1620
QY 1363 -----TGCTTCCACCCCATGGCCATCGAGAACACTCAAGAAAGAAAT 1405
DB 1621 GCCTGGCGGTGAGGCGCGCTTGACTGGGTGTACTACGAGTCAAGGCACATATCCACT 1680
QY 1406 GTCCCTTCTACTA-----GTGGGTAGATTCCTTAGG-----GCCGCGCGA 1449
DB 1681 GCTCGGTGAAGCAGAGAACTCGGTGGCGGCAAAATCGGAGGCTGATTCGCGCGCGCGA 1740
QY 1450 ATTAATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTGGGCACCTTACGAGAAAGACCG 1509
DB 1741 ATTAATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTGGGCACCTTACGAGAAAGACCG 1800
QY 1510 GTCTTGCTAGATTTCTAATCAAGAGAGTGTGAAATGCCATTTGCTCAGAGATGCGAGGCT 1569
DB 1801 GTCTTGCTAGATTTCTAATCAAGAGAGTGTGAAATGCCATTTGCTCAGAGATGCGAGGCT 1860
QY 1570 TCATTTTGTACTTTTTTATTTGTAACTATATAGTATAGGATTTTTTTTGTGCAATTTG 1629
DB 1861 TCATTTTGTACTTTTTTATTTGTAACTATATAGTATAGGATTTTTTTTGTGCAATTTG 1920
QY 1630 TTTCTTCTGACAGGCTGCTCCTGATCAGCCTATCTCGCAGCTGATGAATATCTTGTG 1689
DB 1921 TTTCTTCTGACAGGCTGCTCCTGATCAGCCTATCTCGCAGCTGATGAATATCTTGTG 1980
QY 1690 GTAGGGTTTGGGAAATCAATTCAGTTTGTGATTTTTTCTTGATTTTCCCACTCCTCTT 1749
DB 1981 GTAGGGTTTGGGAAATCAATTCAGTTTGTGATTTTTTCTTGATTTTCCCACTCCTCTT 2040
QY 1750 CAGAGTACAGAAATTAAGTGAAGAGTTGTTGTGCAAGCTTATCGATAAGCTTTAATG 1809
DB 2041 CAGAGTACAGAAATTAAGTGAAGAGTTGTTGTGCAAGCTTATCGATAAGCTTTAATG 2100
QY 1810 CGGTAGTTTATCAGAGTTAAATTTGCTAAACGAGTCAGGACCGGTGATGAATCTAACAA 1869
DB 2101 CGGTAGTTTATCAGAGTTAAATTTGCTAAACGAGTCAGGACCGGTGATGAATCTAACAA 2160
QY 1870 TGCGCTCATCGTCTATCTCGGCACCGTCCACCTGGATGCTGTAGGCAATAGGCTTGGTTAT 1929
DB 2161 TGGCGTCTATGCTATCTCTCGGCACCGTCCACCTGGATGCTGTAGGCAATAGGCTTGGTTAT 2220

QY 1930 GCGGTACTTGCGGGGCTCTTTCGGGATATCGTCCATTTCCGACAGCATCGCCACTACTA 1989
DB 2221 GCGGTACTTGCGGGGCTCTTTCGGGATATCGTCCATTTCCGACAGCATCGCCACTACTA 2280
QY 1990 TGGCGTGTCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCCGTTCTCGGAGC 2049
DB 2281 TGGCGTGTCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCCGTTCTCGGAGC 2340
QY 2050 ACTGTCCGACCGCTTTGGCCGCGCCAGTCTGCTGCTTCGCTTCTGCTACTTGGAGCACTAT 2109
DB 2341 ACTGTCCGACCGCTTTGGCCGCGCCAGTCTGCTGCTTCGCTTCTGCTACTTGGAGCACTAT 2400
QY 2110 CGATPACGGGATCATGCGGACACACCCGTCCTGTGATCTATCGAATCTAAATGTTAAGT 2169
DB 2401 CGACTPACGGGATCATGCGGACACACCCGTCCTGTGATCTATCGAATCTAAATGTTAAGT 2460
QY 2170 TAAATCTCTAAATTAATTAATTAAGTCCAGTTTCTCCATACGAACCTTTACAGCATTCG 2229
DB 2461 TAAATCTCTAAATTAATTAATTAAGTCCAGTTTCTCCATACGAACCTTTACAGCATTCG 2520
QY 2230 GGTGAGCATCTAGACCTTTCAACAGCAGCCAGATCCATCACTGCTTGGCCAAATATGTTCA 2289
DB 2521 GGTGAGCATCTAGACCTTTCAACAGCAGCCAGATCCATCACTGCTTGGCCAAATATGTTCA 2580
QY 2290 GTCCCTCAGGAGTTACGCTTTGTGAAGTGAAGTAACTTCTGGAAGTTGCGATGTTAACTC 2349
DB 2581 GTCCCTCAGGAGTTACGCTTTGTGAAGTGAAGTAACTTCTGGAAGTTGCGATGTTAACTC 2640
QY 2350 CGCTGTAATTGACGGGCATATCCGTACGTTGGCAAGTGTGTTGTTGTTACCGGAGGAGTAAT 2409
DB 2641 CGCTGTAATTGACGGGCATATCCGTACGTTGGCAAGTGTGTTGTTGTTACCGGAGGAGTAAT 2700
QY 2410 CTCACAACCTCTCTGGAGTAGGACCAACAAACACAGATCCAGCGTGTGTTACTTGAT 2469
DB 2701 CTCACAACCTCTCTGGAGTAGGACCAACAAACACAGATCCAGCGTGTGTTACTTGAT 2760
QY 2470 CAACATAAGAAAGCAATTTCTCGATTTTCAGAGTCAAGTGTTCAGGAGCGTACTGATGG 2529
DB 2761 CAACATAAGAAAGCAATTTCTCGATTTTCAGAGTCAAGTGTTCAGGAGCGTACTGATGG 2820
QY 2530 ACATTTCAAAGCCTGCTGTTAGTTGCAACCGATAGGTTGTAGTGTGCAATACACT 2589
DB 2821 ACATTTCAAAGCCTGCTGTTAGTTGCAACCGATAGGTTGTAGTGTGCAATACACT 2880
QY 2590 TCGGTACAAATTTCAACCTTGGCAACTGCACAGCTTGTGTTGTTGAAACAGCATCTTCAATTC 2649
DB 2881 TCGGTACAAATTTCAACCTTGGCAACTGCACAGCTTGTGTTGTTGAAACAGCATCTTCAATTC 2940
QY 2650 TGGCAAGCTCCTGTCTCTCATATCGACAGCAACAGAAATCACCTGGGAATCAATACCAT 2709
DB 2941 TGGCAAGCTCCTGTCTCTCATATCGACAGCAACAGAAATCACCTGGGAATCAATACCAT 3000
QY 2710 GTTCAGCTTGAGACAGAGAGTCTGAGGCAACGAATCTGGATCAGCGTATTTATCAGCAA 2769
DB 3001 GTTCAGCTTGAG-CAGAAAGTCTGAGGCAACGAATCTGGATCAGCGTATTTATCAGCAA 3059
QY 2770 TAACTAGAACTTCAGAAAGCCGAGCAGGCAATGTCAATACTACAGGCGCTGATGTGAT 2829
DB 3060 TAACTAGAACTTCAGAAAGCCGAGCAGGCAATGTCAATACTACAGGCGCTGATGTGAT 3119
QY 2830 TTTGAAACCATCATCTTGGCAGCAGTAAAGAACTGGTTTCTTGGACCAATATTTGTCAC 2889
DB 3120 TTTGAAACCATCATCTTGGCAGCAGTAAAGAACTGGTTTCTTGGACCAATATTTGTCAC 3179
QY 2890 ACTTAGGAACAGTTTCTGTTCCGTAAGCCATAGCAGCTACTGCTTGGGCGCTCCTGCTA 2949
DB 3180 ACTTAGGAACAGTTTCTGTTCCGTAAGCCATAGCAGCTACTGCTTGGGCGCTCCTGCTA 3239
QY 2950 GCAAGATACATTTAGACCAACCTTGTGGGCAACGATAGTACTTCTGGGAGTAAAGGTAC 3009
DB 3240 GCAAGATACATTTAGACCAACCTTGTGGGCAACGATAGTACTTCTGGGAGTAAAGGTAC 3299
QY 3010 CATCTCTTCTTAGGTGGAGATGCAAAAAACAATTTCTTTCACACCAAGCAACTTTTGGCAGAA 3069

QY 4998 ----- 4997
Db 5519 GGCATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCGATGCCGAGAGTTGTTTC 5578
QY 4998 ----- 4997
Db 5579 TGAACATGCGAAAGTAGCGTTGCCAATGAATGTACAGATGAGATGGTTCAGACTAAACT 5638
QY 4998 ----- 4997
Db 5639 GGCTGACGGAAATTTATGCTCTCCGACCATCAAGCAATTTATCCGTACTCTCTGATGATG 5698
QY 4998 ----- 4997
Db 5699 CATGTTACTCACCCTCGCATCCCGGAAACAGCANTCCAGGTATTAGAAGATATC 5758
QY 4998 ----- 4997
Db 5759 CTGATTCAGGTGAAATAATTGTTGATCGCTGCGCAGTGTTCCTGCGCGGTTGCATTCGA 5818
QY 4998 ----- 4997
Db 5819 TTCCTGTTGTAATGTCCTTTTAAACAGCGATCCGCTATTTGCTCTCGCTCAGCGCGCAAT 5878
QY 4998 ----- 4997
Db 5879 CACGAATGAATAACGGTTTGGTTGATGCGAGTGAATTTTGATGACGAGCGTAATGGCTGGC 5938
QY 4998 ----- 4997
Db 5939 CTGTTGAACAAGTCTGGAAGAAATGCAATAGCTTTTGCCATTTCTACCGGATTCAGTCG 5998
QY 4998 ----- 4997
Db 5999 TCACTCATGGTGATTTCTCACTTGATAACCTTATTTTGGACGAGGGAATTAATAGGTT 6058
QY 4998 ----- 4997
Db 6059 GTATTGATGTTGGACGAGTCGGAATCGACAGCCGATACCAGGATCTTGCCATCCTATGGA 6118
QY 4998 ----- 4997
Db 6119 ACTGCTCGGTGAGTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATGTTATTG 6178
QY 4998 ----- 4997
Db 6179 ATAATCCTGATATGAATAAATTCGAGTTTCATTTGATGCTCGATGAGTTTTCATATCAG 6238
QY 4998 ----- 4997
Db 6239 AATTGGTTAATGTTGTAACACTGGCAGACATTACGCTGACTTGACGGGACGGCGGT 6298
QY 4998 ----- 4997
Db 6299 TTGTTGAATAAATCGAACTTTTGTGCTGAGTTGAAGGATCAGATCAGCATCTTCCCGCAA 6358
QY 4998 ----- 4997
Db 6359 CGCAGACCGTTCCGTGGCAAGCAAAAGTTCAAAATCACCNACTGCTCCACCTACACAA 6418
QY 4998 ----- 4997
Db 6419 AGCTCTCATCAACCGTGCCTCCCTCACTTTCTGCTGGATGATGGGCGGATTCAGGCGCTG 6478
QY 4998 ----- 4997
Db 6479 GTATGATCAGCAACACTTTTTCAGGAGGAGACCTCAGCGCGGAGACCTCAGCGGCTGCGG 6538
QY 5009 CTTTGGCGCGGCTGCTCAACGGCCTCAACTACTACTGGGCTGCTTCTTAATGACGGA 5068
Db 6539 TCCACGCGCGGCTGCTCAACGGCCTCAACTACTACTGGGCTGCTTCTTAATGACGGA 6598

QY 5069 CTCGCATAAGGAGAGCGTCCAGTATCTATGATTGGAAGTATGGAAATGTTGATACCCGC 5128
Db 6599 GTCCGCATTAAGGAGAGCGTCCAGTATCTATGATTGGAAGTATGGAAATGTTGATACCCGC 6658
QY 5129 ATTCTTCAGTCTCTGAGTCTCTCTATCAGATTATGCCCACTAAAGCAACCGGAGGAGG 5188
Db 6659 ATTCTTCAGTCTCTGAGTCTCTCTATCAGATTATGCCCACTAAAGCAACCGGAGGAGG 6718
QY 5189 AGATTTCATGTTAAATTTCTCTGACTTTTGTCTATCAGTAGACTCGAACTGTGAGACTAT 5248
Db 6719 AGATTTCATGTTAAATTTCTCTGACTTTTGTCTATCAGTAGACTCGAACTGTGAGACTAT 6778
QY 5249 CTCGGTTATGACAGCAGAAATGTCCTCTTGGAGACAGTAAATGAATCCACCACCAATAA 5308
Db 6779 CTCGGTTATGACAGCAGAAATGTCCTCTTGGAGACAGTAAATGAATCCACCACCAATAA 6838
QY 5309 GAAATCCTTGTATCAGGAACAAACTTCTTGTTCGAACTTTTTCGGTGCCTTGAACATAT 5368
Db 6839 GAAATCCTTGTATCAGGAACAAACTTCTTGTTCGAACTTTTTCGGTGCCTTGAACATAT 6898
QY 5369 AAAATGTAGAGTGAATATGTCGGTAGGAATGGAAGCGGCGGAAATGCTTACCTTCGACC 5428
Db 6899 AAAATGTAGAGTGAATATGTCGGTAGGAATGGAAGCGGCGGAAATGCTTACCTTCGACC 6958
QY 5429 TTCAAGAGCTATGAGGTTTGTAGATCTCATCTGCAACTTCAGTGACAAGCTTGTCTATT 5488
Db 6959 TTCAAGAGCTATGAGGTTTGTAGATCTCATCTGCAACTTCAGTGACAAGCTTGTCTATT 7018
QY 5489 TCGTTCAAAACCTCCGAAATCCAGAGAAATCAAAAGTTGTTGCTACTATTGATCCAAAGC 5548
Db 7019 TCGTTCAAAACCTCCGAAATCCAGAGAAATCAAAAGTTGTTGCTACTATTGATCCAAAGC 7078
QY 5549 CAGTCCGCTCTGAAACTGACAATAGTCTGCTGCTGTTTGGAGTCTATCTTTGATGAAT 5608
Db 7079 CAGTCCGCTCTGAAACTGACAATAGTCTGCTGCTGTTTGGAGTCTATCTTTGATGAAT 7138
QY 5609 AAATCTAGTCTTGAATCTAAATTAATCTTTCAGGAGCCAAAGCGGATAAATACCCAAATCTAA 5668
Db 7139 AAATCTAGTCTTGAATCTAAATTAATCTTTCAGGAGCCAAAGCGGATAAATACCCAAATCTAA 7198
QY 5669 AACTCTTTTAAACGTTAAAGGACAAGTATGTCCTGTATTAAGACCCCAATCAGCT 5728
Db 7199 AACTCTTTTAAACGTTAAAGGACAAGTATGTCCTGTATTAAGACCCCAATCAGCT 7258
QY 5729 CGTAGTCTGATCTCATCAACTTCAGGAGCCACTATCTTGTGTTTAGAGAAATTTTCGGAGA 5788
Db 7259 CGTAGTCTGATCTCATCAACTTCAGGAGCCACTATCTTGTGTTTAGAGAAATTTTCGGAGA 7318
QY 5789 TCGATATCGAGAAATAAGGTACGCTGATTTTAAACGTTAAATTTATCTCAAGATCTCTGC 5848
Db 7319 TCGATATCGAGAAATAAGGTACGCTGATTTTAAACGTTAAATTTATCTCAAGATCTCTGC 7378
QY 5849 CTCGCGCGTTTCGGTGATGACGTTGAAACCTCTGACATGCACTCCCGGAGCGGCTC 5908
Db 7379 CTCGCGCGTTTCGGTGATGACGTTGAAACCTCTGACATGCACTCCCGGAGCGGCTC 7438
QY 5909 ACAGCTTGTCTGTAAGCGGATCCCGGAGCAGACAAGCCCTCAGGCGGCTCAGCGGCT 5968
Db 7439 ACAGCTTGTCTGTAAGCGGATCCCGGAGCAGACAAGCCCTCAGGCGGCTCAGCGGCT 7498
QY 5969 GTTGGCGGCTTCGGGCGCAGCAGTACCCAGTACGTTAGCGAGTGTATCTACT 6028
Db 7499 GTTGGCGGCTTCGGGCGCAGCAGTACCCAGTACGTTAGCGAGTGTATCTACT 7558
QY 6029 GGCCTTAACTATTCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAA 6088
Db 7559 GGCCTTAACTATTCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAA 7618
QY 6089 TACCGCAGATGCTGTAAGGAGAAATACCGCATCAGCGCTCTTCCGCTTCCTCGCTCA 6148
Db 7619 TACCGCAGATGCTGTAAGGAGAAATACCGCATCAGCGCTCTTCCGCTTCCTCGCTCA 7678
QY 6149 CTGACTCGCTCGGCTCGGCTTCGGCTGCGGCGGAGCGGATATCAGCTCACTCAAGGCGG 6208

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2004, 11:31:21 ; Search time 20392.8 Seconds
(without alignments)
17515.470 Million cell updates/sec

Title: US-10-030-390-3
Perfect score: 8241
Sequence: 1 agatctaacatccaagacg.....aaaataacagttattattatg 8241

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.hcg.*
- 3: gb.in.*
- 4: gb.ov.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.scs.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.or.*
- 21: em.ov.*
- 22: em.pat.*
- 23: em.ph.*
- 24: em.pl.*
- 25: em.ro.*
- 26: em.sts.*
- 27: em.un.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.hcg.hum.*
- 31: em.hcg.inv.*
- 32: em.hcg.other.*
- 33: em.hcg.mus.*
- 34: em.hcg.pln.*
- 35: em.hcg.rcd.*
- 36: em.hcg.mam.*
- 37: em.hcg.vrt.*
- 38: em.sy.*
- 39: em.hgo.hum.*
- 40: em.hgo.mus.*
- 41: em.hgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	8241	100.0	8241	6	AX069291	AX069291 Sequence
2	7555.6	91.7	8598	6	E16036	E16036 cDNA encodi
3	7223.6	87.7	7925	12	AX178045	AX178045 Expressio
4	6726.6	81.6	8020	8	PPPIC9	Z46233 P.pastoris
5	6713.6	81.5	7757	8	PPPIC3	Z46231 P.pastoris
6	6687.6	81.2	9408	6	AX156465	AX156465 Sequence
7	6561.6	79.6	9276	6	AX149597	AX149597 Sequence
8	6561.6	79.6	9276	6	AX207941	AX207941 Sequence
9	6383.8	77.5	9776	6	AX146618	AX146618 Sequence
10	6383.8	77.5	9776	6	AX417106	AX417106 Sequence
11	6250.2	75.8	9139	12	AX178634	AX178634 Expressio
12	5709.8	69.3	10462	6	AX146621	AX146621 Sequence
13	5677.8	68.9	10462	6	AX417109	AX417109 Sequence
14	5677.8	68.9	10492	6	AX146619	AX146619 Sequence
15	5665.8	68.8	10491	6	AX417107	AX417107 Sequence
16	5657.8	68.7	10512	6	AX146620	AX146620 Sequence
17	5657.8	68.7	10512	6	AX417108	AX417108 Sequence
18	5509.6	66.9	9273	8	PPPIC9K	Z46234 P.pastoris
19	5496.6	66.7	9010	8	PPPIC3K	Z46232 P.pastoris
20	5142.8	62.4	8584	6	I63561	I63561 Sequence 66
21	5130.8	62.3	8590	6	I63563	I63563 Sequence 70
22	4565.2	55.4	8157	6	I63560	I63560 Sequence 65
C 23	2659	32.3	2659	8	ARI06448	ARI06448 Sequence
C 24	2634.2	32.0	2659	8	PPHIS4G	X56180 P.pastoris
C 25	2634.2	32.0	5270	12	AF027958	AF027958 Expressio
C 26	2634.2	32.0	5546	12	AF027959	AF027959 Expressio
C 27	2634.2	32.0	5565	12	AF027960	AF027960 Expressio
C 28	2634.2	32.0	5796	12	AF027961	AF027961 Expressio
C 29	2627.4	31.9	5839	12	CVU14122	U14122 Cloning vec
C 30	2621.4	31.8	5239	12	CVU14124	U14124 Cloning vec
C 31	2621.4	31.8	5284	12	CVU14123	U14123 Cloning vec
C 32	2621.4	31.8	5508	12	CVU14127	U14127 Cloning vec
C 33	2621.4	31.8	5535	12	CVU14117	U14117 Cloning vec
C 34	2621.4	31.8	5619	12	CVU14119	U14119 Cloning vec
C 35	2621.4	31.8	5821	12	CVU14118	U14118 Cloning vec
C 36	2621.4	31.8	5822	12	CVU14121	U14121 Cloning vec
C 37	2621.4	31.8	5634	12	CVU14120	U14120 Cloning vec
C 38	2621.4	31.8	5634	12	CVU14125	U14125 Cloning vec
C 39	2621.4	31.8	5936	12	CVU14116	U14116 Cloning vec
C 40	2620.2	31.8	2652	8	PPU14126	U14126 Pichia past
C 41	2325.8	28.2	3753	12	SYNTRPC	M33624 Cloning vec
C 42	2325.8	28.2	3763	12	SYNTRPA	M33622 Cloning vec
C 43	2325.8	28.2	3779	12	SYNTRPH	M32985 Expression
C 44	2324.2	28.2	3771	12	SYNTRPB	M33623 Cloning vec
C 45	2324.2	28.2	3772	12	SYNTRPD	M33625 Cloning vec

ALIGNMENTS

RESULT 1	AX069291	Sequence 3 from Patent WO0102570.	8241 bp	DNA	linear	PAT 25-JAN-2001
LOCUS	AX069291					
DEFINITION	AX069291					
ACCESSION	AX069291					
VERSION	AX069291.1	GI:12579164				
KEYWORDS						
SOURCE		synthetic construct				
ORGANISM		synthetic construct				
		artificial sequences.				
REFERENCE	1					
AUTHORS		Hans,W.C., Steidler,J. and Remaut,E.R.				
TITLE		Delivery of trefol peptides				
JOURNAL		Patent: WO 0102570-A 3 11-JAN-2001;				
		Vlaams Interuniversitair Instituut voor Biotechnologie (BE)				

FEATURES		Location/Qualifiers	
source		1..8241	
		/organism="synthetic construct"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:32630"	
		/note="Synthetic construct : plasmid"	
ORIGIN			
Query Match		100.0%; Score 8241; DB 6; Length 8241;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 8241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	AGATCTAACATCCAAAGACGAAAGGTTGAATGAAACCTTTTGGCCATCCGACATCCACAG	60
DB	1	AGATCTAACATCCAAAGACGAAAGGTTGAATGAAACCTTTTGGCCATCCGACATCCACAG	60
QY	61	GTCCATTCTCACATAGTCCCAACGCAACAGGAGGGATACATAGCAGCAGACCGT	120
DB	61	GTCCATTCTCACATAGTCCCAACGCAACAGGAGGGATACATAGCAGCAGACCGT	120
QY	121	TGCAAAACGACGACCTCCACTCTCTCTCTCAACACCCACTTTTGCCATCGAAAAACC	180
DB	121	TGCAAAACGACGACCTCCACTCTCTCTCTCAACACCCACTTTTGCCATCGAAAAACC	180
QY	181	AGCCAGTATTGGGCTGTGATGGAGCTCGTCAATCCAAATCCTTCTATTAGGCTACTA	240
DB	181	AGCCAGTATTGGGCTGTGATGGAGCTCGTCAATCCAAATCCTTCTATTAGGCTACTA	240
QY	241	ACACATGACTTTATTAGCTGTCTATCTCTGCGCCCTCGCGAGGTTCTATGTTGTTTA	300
DB	241	ACACATGACTTTATTAGCTGTCTATCTCTGCGCCCTCGCGAGGTTCTATGTTGTTTA	300
QY	301	TTTCCGAATGCAACAGCTCCGCAATTACCCGCAACATCACTCCAGATGAGGCTTCTG	360
DB	301	TTTCCGAATGCAACAGCTCCGCAATTACCCGCAACATCACTCCAGATGAGGCTTCTG	360
QY	361	AGTGTGGGTCAAAATAGTTTCATGTTCCCAATGGCCCAAACTGACATTTAAAGCCT	420
DB	361	AGTGTGGGTCAAAATAGTTTCATGTTCCCAATGGCCCAAACTGACATTTAAAGCCT	420
QY	421	GCTTTGGAACCTAATATGACAAAAGCGTGATCTCATCCAAAGTGAACCTAAGTTGGTTGG	480
DB	421	GCTTTGGAACCTAATATGACAAAAGCGTGATCTCATCCAAAGTGAACCTAAGTTGGTTGG	480
QY	481	TTGAAATGCTAACCGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCCGCAATACCGTTTGT	540
DB	481	TTGAAATGCTAACCGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCCGCAATACCGTTTGT	540
QY	541	CTTGTGTTGATATGATTGACGAATGCTCAAAAATPAATCTCAATTAAGTCTTAGCGCAGTCT	600
DB	541	CTTGTGTTGATATGATTGACGAATGCTCAAAAATPAATCTCAATTAAGTCTTAGCGCAGTCT	600
QY	601	CTCTATCGCTTCTGAACCCCGGTGCACCTGTGCGGAAACCAATGCGGAAAACACCCCGCT	660
DB	601	CTCTATCGCTTCTGAACCCCGGTGCACCTGTGCGGAAACCAATGCGGAAAACACCCCGCT	660
QY	661	TTTGGATGATATGATGCTCCACATGTTATGCTTCCAAAGTCTGCTGGTGGGATACT	720
DB	661	TTTGGATGATATGATGCTCCACATGTTATGCTTCCAAAGTCTGCTGGTGGGATACT	720
QY	721	GCTGATAGCCTTAACGTTTCATGATCAAAATTTAACTGTTCTAAACCCCTACTTGACAGCAAT	780
DB	721	GCTGATAGCCTTAACGTTTCATGATCAAAATTTAACTGTTCTAAACCCCTACTTGACAGCAAT	780
QY	781	ATATAACAGAGAAGCTGCCCTGCTTAAACCTTTTATGATCATGATTAATAGCTT	840
DB	781	ATATAACAGAGAAGCTGCCCTGCTTAAACCTTTTATGATCATGATTAATAGCTT	840
QY	841	ACTTTTCAATTTGGACCTGGTTCCAAATTTGACAGCTTTTGAATTTAAACGACTTTTAACGA	900
DB	841	ACTTTTCAATTTGGACCTGGTTCCAAATTTGACAGCTTTTGAATTTAAACGACTTTTAACGA	900
QY	901	CAACTTGAGAGATCAAAAACAACCTAATTAATTCGAAGGATCCAAACGATGAGATTTCCT	960
DB	901	CAACTTGAGAGATCAAAAACAACCTAATTAATTCGAAGGATCCAAACGATGAGATTTCCT	960
QY	961	TCAATTTTACTGCAAGTTTATTCGAGCATCTCCGATAGCTGCTCCAGTCAACACT	1020
DB	961	TCAATTTTACTGCAAGTTTATTCGAGCATCTCCGATAGCTGCTCCAGTCAACACT	1020
QY	1021	CAACACAGAAAGATGAAACGGCACAAATTCGGCTGAAGCTGTCACTCGGTTACTCAGATTTA	1080
DB	1021	CAACACAGAAAGATGAAACGGCACAAATTCGGCTGAAGCTGTCACTCGGTTACTCAGATTTA	1080
QY	1081	GAAAGGATTTTCGATGTTGCTGTTTCCATATTTTCCAAACAGCACAAATAACGGGTTATTG	1140
DB	1081	GAAAGGATTTTCGATGTTGCTGTTTCCATATTTTCCAAACAGCACAAATAACGGGTTATTG	1140
QY	1141	TTTATAAATACTACTATTTCAGCAATTCCTGCTAAAGAAAGAGGGTATCTCTCGAGAAA	1200
DB	1141	TTTATAAATACTACTATTTCAGCAATTCCTGCTAAAGAAAGAGGGTATCTCTCGAGAAA	1200
QY	1201	AGAGAGGCTGAAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG	1260
DB	1201	AGAGAGGCTGAAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG	1260
QY	1261	GCCCGCCGGAGAGGATAAATTTGGGCTTCCCGGTGTCAACGCGCAGCAGTGCCAGGAG	1320
DB	1261	GCCCGCCGGAGAGGATAAATTTGGGCTTCCCGGTGTCAACGCGCAGCAGTGCCAGGAG	1320
QY	1321	AGAGGTTCTGTTTGTATGACAGTGTCCGGGATTTCCGCTGCTCTCCAGCCCATGGCC	1380
DB	1321	AGAGGTTCTGTTTGTATGACAGTGTCCGGGATTTCCGCTGCTCTCCAGCCCATGGCC	1380
QY	1381	ATCGAGAACTCAAGAAAGAAATGTCCTTCTAAGTGTGGGTAGAAATTCCTTAGGG	1440
DB	1381	ATCGAGAACTCAAGAAAGAAATGTCCTTCTAAGTGTGGGTAGAAATTCCTTAGGG	1440
QY	1441	CGGCGCCCAATTAATTCGCTTAGCATGACTGTTCTCAGTTCAGTGTGGGCACTTACG	1500
DB	1441	CGGCGCCCAATTAATTCGCTTAGCATGACTGTTCTCAGTTCAGTGTGGGCACTTACG	1500
QY	1501	AGAAGACCGCTTGTCTAGATTCTTAACAAGAGATGTCAAGATGCCATTTCCTAGAG	1560
DB	1501	AGAAGACCGCTTGTCTAGATTCTTAACAAGAGATGTCAAGATGCCATTTCCTAGAG	1560
QY	1561	ATCGAGGCTTCATTTTGTATCTTTTATTTGTAACCTATATATAGTATAGGATTTTTTT	1620
DB	1561	ATCGAGGCTTCATTTTGTATCTTTTATTTGTAACCTATATATAGTATAGGATTTTTTT	1620
QY	1621	GTCAATTTTGTCTTCTCGTAGAGCTTGTCTCTGATCAGCCTATCTCGAGCTGATGAA	1680
DB	1621	GTCAATTTTGTCTTCTCGTAGAGCTTGTCTCTGATCAGCCTATCTCGAGCTGATGAA	1680
QY	1681	TATCTTGTGTAGGGTTTGGGAAATCATTCGAGTTTGAATTTTCTTGTGTTATTTCCC	1740
DB	1681	TATCTTGTGTAGGGTTTGGGAAATCATTCGAGTTTGAATTTTCTTGTGTTATTTCCC	1740
QY	1741	ACTCTCTTTCAGAGTACAGAAAGATTAAAGTGAAGAGTTTGTGCAAGCTTATCGATAA	1800
DB	1741	ACTCTCTTTCAGAGTACAGAAAGATTAAAGTGAAGAGTTTGTGCAAGCTTATCGATAA	1800
QY	1801	GCTTTAATTCGGTAGTTTATCACAGTTAAATGCTAACGAGTCAAGCAGCTGTATGAA	1860
DB	1801	GCTTTAATTCGGTAGTTTATCACAGTTAAATGCTAACGAGTCAAGCAGCTGTATGAA	1860
QY	1861	ATCTAAACATGCGCTCATCTCCGCAACCGTCAACCTGGATGCTGTAGGCATAGG	1920
DB	1861	ATCTAAACATGCGCTCATCTCCGCAACCGTCAACCTGGATGCTGTAGGCATAGG	1920
QY	1921	CTTGTGTATGCGGTACTGCGGGGCTTTCGGGGATATCGTCCATTCGACAGCATCGC	1980
DB	1921	CTTGTGTATGCGGTACTGCGGGGCTTTCGGGGATATCGTCCATTCGACAGCATCGC	1980
QY	1981	CAGTCACATATGCGGTGCTGCTAGCGCTATATGCTGATGCAATTTCTATCGCACCCGT	2040
DB	1981	CAGTCACATATGCGGTGCTGCTAGCGCTATATGCTGATGCAATTTCTATCGCACCCGT	2040

1981 CAGTCACTATGGCGTGTCTAGCGCTATATGGTTGATGCAATTTCTATGCGACCCCGT 2040 Db
2041 TCTCGGAGCATGTCGCGACCGCTTTGGCGCGCCAGTCTCTGCTCGCTTCGTTACTTGG 2100 Qy
2041 TCTCGGAGCATGTCGCGACCGCTTTGGCGCGCCAGTCTCTGCTCGCTTCGTTACTTGG 2100 Db
2101 AGCCACTATCGACTACGCGATCATGGCGACACACCGCGTCTCTGGATCTATCGAATCTA 2160 Qy
2101 AGCCACTATCGACTACGCGATCATGGCGACACACCGCGTCTCTGGATCTATCGAATCTA 2160 Db
2161 AATGTAGTTAAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2220 Qy
2161 AATGTAGTTAAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2220 Db
2221 CAGCATTGCGGTGAGCATCTAGACCTTCAACAGCAGCCAGATCCATCACTGCTTGGCCAA 2280 Qy
2221 CAGCATTGCGGTGAGCATCTAGACCTTCAACAGCAGCCAGATCCATCACTGCTTGGCCAA 2280 Db
2281 TATGTTTCAGTCCCTCAGGAGTTACGTTCTGTAAGTGAATCTTCTGGAAGTTTCAG 2340 Qy
2281 TATGTTTCAGTCCCTCAGGAGTTACGTTCTGTAAGTGAATCTTCTGGAAGTTTCAG 2340 Db
2341 TGTAACTCCGCTGTATTGACGGGCATATCCGTACGTTGGCAAGTGTGGTACCGG 2400 Qy
2341 TGTAACTCCGCTGTATTGACGGGCATATCCGTACGTTGGCAAGTGTGGTACCGG 2400 Db
2401 AGGAGTAATCTCCAACTCTCTGAGAGTAGGACCAACAAACAGATCCAGCGTGT 2460 Qy
2401 AGGAGTAATCTCCAACTCTCTGAGAGTAGGACCAACAAACAGATCCAGCGTGT 2460 Db
2461 GTACTGATCAACATTAAGAAGCATCTCTGATTGAGGATCAAGTGTTCAGGAGCGT 2520 Qy
2461 GTACTGATCAACATTAAGAAGCATCTCTGATTGAGGATCAAGTGTTCAGGAGCGT 2520 Db
2521 ACTGATTGGACATTTCCAAAGCTGCTGTAAGTTGCAACCGATAGGTTGTAGAGTGT 2580 Qy
2521 ACTGATTGGACATTTCCAAAGCTGCTGTAAGTTGCAACCGATAGGTTGTAGAGTGT 2580 Db
2581 CAATACACTTGGTACAAATTTCAACCTTGGCACTGACAGCTTGGTTGACAGCAT 2640 Qy
2581 CAATACACTTGGTACAAATTTCAACCTTGGCACTGACAGCTTGGTTGACAGCAT 2640 Db
2641 CTTCAATTTCTGGCAAGCTCTTGTCTGTCATATCGACAGCCAAACAGATCACTGGGAAT 2700 Qy
2641 CTTCAATTTCTGGCAAGCTCTTGTCTGTCATATCGACAGCCAAACAGATCACTGGGAAT 2700 Db
2701 CAATACCAATGTTGACGACAGAGTCTGAGGCAACGAATCTGGATCAGCGTATT 2760 Qy
2701 CAATACCAATGTTGACGACAGAGTCTGAGGCAACGAATCTGGATCAGCGTATT 2760 Db
2761 TATCAGCAATTAAGTAACTTCAAGAGCCAGAGGATGTCAATCTACACAGGCGTG 2820 Qy
2761 TATCAGCAATTAAGTAACTTCAAGAGCCAGAGGATGTCAATCTACACAGGCGTG 2820 Db
2821 ATGTGTCATTTTGAACCATCATCTTTGGCAGCAGTAACGAACTGGTTCTTGACCAATA 2880 Qy
2821 ATGTGTCATTTTGAACCATCATCTTTGGCAGCAGTAACGAACTGGTTCTTGACCAATA 2880 Db
2881 TTTTGTACACTTAGGAAGATTTCTGTCGATAGCCATAGCAGTACTGCTGGGCG 2940 Qy
2881 TTTTGTACACTTAGGAAGATTTCTGTCGATAGCCATAGCAGTACTGCTGGGCG 2940 Db
2941 CTCCTGCTAGCAGTACATCTTAGCAACCACTTGTGGGCAACGATGATGATCTTGGG 3000 Qy
2941 CTCCTGCTAGCAGTACATCTTAGCAACCACTTGTGGGCAACGATGATGATCTTGGG 3000 Db
3001 TAAGGTACATCTTCTTAGGTGGAGTGCAGAAACAAATTTCTTGGCAACGAACTT 3060 Qy
3001 TAAGGTACATCTTCTTAGGTGGAGTGCAGAAACAAATTTCTTGGCAACGAACTT 3060 Db
3061 TGGCAGGAACCCAGCATCAGGAGTGCAGGAGGATTCGGTTCCACGAGGAATAT 3120 Qy
3061 TGGCAGGAACCCAGCATCAGGAGTGCAGGAGGATTCGGTTCCACGAGGAATAT 3120 Db

3121 AGAGGCCAACTTTCTCAATAGGTCTTGCAAAACGAGAGGAGACTACACAGGGCAAGTCT 3180 Qy
3121 AGAGGCCAACTTTCTCAATAGGTCTTGCAAAACGAGAGGAGACTACACAGGGCAAGTCT 3180 Db
3181 CAACTTGAACGCTCTCGTTAGTTGAGCTTCATGAAATTTCTCTGAGCTTATCTATAGAGA 3240 Qy
3181 CAACTTGAACGCTCTCGTTAGTTGAGCTTCATGAAATTTCTCTGAGCTTATCTATAGAGA 3240 Db
3241 GATCAATGCTCTCTTAACGTTATCTGGCAATTTGATAGTTCTCTGGGAAAGAGCTT 3300 Qy
3241 GATCAATGCTCTCTTAACGTTATCTGGCAATTTGATAGTTCTCTGGGAAAGAGCTT 3300 Db
3301 CTAACACAGCTCTCTTAACGTTATCTGGCAATTTGATAGTTCTCTGGGAAAGAGCTT 3360 Qy
3301 CTAACACAGCTCTCTTAACGTTATCTGGCAATTTGATAGTTCTCTGGGAAAGAGCTT 3360 Db
3361 TGTCAACATTTTGAAGCAATTTGATAGTTCTCTGGCAATTTGATAGTTCTCTGGGAAAGAGCTT 3420 Qy
3361 TGTCAACATTTTGAAGCAATTTGATAGTTCTCTGGCAATTTGATAGTTCTCTGGGAAAGAGCTT 3420 Db
3421 TTTTCTGGATAGGACGACGAGGAGCTTCAATTAATTTGATAGTTCTCTGGGAAAGAGCTT 3480 Qy
3421 TTTTCTGGATAGGACGACGAGGAGCTTCAATTAATTTGATAGTTCTCTGGGAAAGAGCTT 3480 Db
3481 CAAATTTGACCAATTTCAATTAATTTGATAGTTCTCTGGGAAAGAGCTT 3540 Qy
3481 CAAATTTGACCAATTTCAATTAATTTGATAGTTCTCTGGGAAAGAGCTT 3540 Db
3541 TAGGTTGTTCTTGGTGTATCTGGTGTGATCTCTGGTGTGATCTCTGGTGTGATCTCTGG 3600 Qy
3541 TAGGTTGTTCTTGGTGTATCTGGTGTGATCTCTGGTGTGATCTCTGGTGTGATCTCTGG 3600 Db
3601 ACTTCATATCCAGGTTTCTCTCCACCTGTCACACGTCACACGTCACACGTCACACGTCAC 3660 Qy
3601 ACTTCATATCCAGGTTTCTCTCCACCTGTCACACGTCACACGTCACACGTCACACGTCAC 3660 Db
3661 CTAATGCAAAATAAATAAAGTCAAGCAATTTCCAGGCTATATCTCTCTTGGATTTAGCTT 3720 Qy
3661 CTAATGCAAAATAAATAAAGTCAAGCAATTTCCAGGCTATATCTCTCTTGGATTTAGCTT 3720 Db
3721 CTCGAAAGTCAATCAGCTTCT 3780 Qy
3721 CTCGAAAGTCAATCAGCTTCT 3780 Db
3781 ACCGTTTGGTATAGAAACCTTCTGGAGCATTTCTTACGATCCCAACAGGTTGGCTTCCA 3840 Qy
3781 ACCGTTTGGTATAGAAACCTTCTGGAGCATTTCTTACGATCCCAACAGGTTGGCTTCCA 3840 Db
3841 TGGCTCTAAGACCCCTTTGATTTGGCCAAAACAGGAAGTGGCTTCCAAGTGCAGAAACCAA 3900 Qy
3841 TGGCTCTAAGACCCCTTTGATTTGGCCAAAACAGGAAGTGGCTTCCAAGTGCAGAAACCAA 3900 Db
3901 CACCTGTTTCTCAACCAAAATTTCAAGCAGTCTCCATCACAATCCAATTCGATACCCA 3960 Qy
3901 CACCTGTTTCTCAACCAAAATTTCAAGCAGTCTCCATCACAATCCAATTCGATACCCA 3960 Db
3961 GCAACTTTTGGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 4020 Qy
3961 GCAACTTTTGGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 4020 Db
4021 GGTAGACTCAGTTTGTGCTTATAGCTTCCGGAATAGACTTTTGGACGAGTACACCA 4080 Qy
4021 GGTAGACTCAGTTTGTGCTTATAGCTTCCGGAATAGACTTTTGGACGAGTACACCA 4080 Db
4081 GGGCCCAACGAGTAATTTAGAGAGTACGACCAACCAAGTAGTGAATAGACCATCGGGCGGT 4140 Qy
4081 GGGCCCAACGAGTAATTTAGAGAGTACGACCAACCAAGTAGTGAATAGACCATCGGGCGGT 4140 Db
4141 CAGTAGTCAAGAGCGCAACAAATTTCTACTGACAGGAACTTTTGGACATCTTCAGAA 4200 Qy
4141 CAGTAGTCAAGAGCGCAACAAATTTCTACTGACAGGAACTTTTGGACATCTTCAGAA 4200 Db

Qy	4201	GTTTCGTATTTCAGTAGTCAATTTCCGAGATCAATAATATGGGGATTATATACCAAGAACAAACAG	4260
Db	4201	GTTTCGTATTTCAGTAGTCAATTTCCGAGATCAATAATATGGGGATTATATACCAAGAACAAACAG	4260
Qy	4261	TGGAAGTCACATCTACCAACTTTTCGGCTCTCAGAAAAAGCATAAACAGTTCTCTACTACCGC	4320
Db	4261	TGGAAGTCACATCTACCAACTTTTCGGCTCTCAGAAAAAGCATAAACAGTTCTCTACTACCGC	4320
Qy	4321	CATTAGTCAAACTTTTCAAAATCGCCAGTGGAGAGAAAAAGGCACAGCGATACTAGCAT	4380
Db	4321	CATTAGTCAAACTTTTCAAAATCGCCAGTGGAGAGAAAAAGGCACAGCGATACTAGCAT	4380
Qy	4381	TAGCGGCGAAGGATGCAACTTTATCAACCAAGGTCCTATAGATAAACCTTAGCGCCTGGGA	4440
Db	4381	TAGCGGCGAAGGATGCAAACTTTATCAACCAAGGTCCTATAGATAAACCTTAGCGCCTGGGA	4440
Qy	4441	TCATCCTTTTGGACAACCTTTTCTGCCAAATCTAGGTCCAAAATCACTTTCAATTGATACCAT	4500
Db	4441	TCATCCTTTTGGACAACCTTTTCTGCCAAATCTAGGTCCAAAATCACTTTGATGATACCAT	4500
Qy	4501	TATTGTACAACTTGAGCAAGTTGTGCATCAGCTCCTCAAAATGGCTCTCTGTAAACGGATG	4560
Db	4501	TATTGTACAACTTGAGCAAGTTGTGCATCAGCTCCTCAAAATGGCTCTCTGTAAACGGATG	4560
Qy	4561	ACTCAACTTGCACATAAATTTGAAGCTCAGTCGATTTGAGTGAACCTTGATCAGGTTGTGCA	4620
Db	4561	ACTCAACTTGCACATAAATTTGAAGCTCAGTCGATTTGAGTGAACCTTGATCAGGTTGTGCA	4620
Qy	4621	GCTGGTCAGCAGCATAGGGAACACGCGCTTTTCTCAACAACTCAAGGAATTTACAACT	4680
Db	4621	GCTGGTCAGCAGCATAGGGAACACGCGCTTTTCTCAACAACTCAAGGAATTTACAACT	4680
Qy	4681	CTGCACACACTTGCCTATGCGAGTAGCAAGGAATGTCATACTTGAAGTCGGACAGTGAG	4740
Db	4681	CTGCACACACTTGCCTATGCGAGTAGCAAGGAATGTCATACTTGAAGTCGGACAGTGAG	4740
Qy	4741	TGTAGTCTTGAGAAATTTCTGAAGCCGTAATTTTATTATCAGTGAGTCAGTCATCAGGAGA	4800
Db	4741	TGTAGTCTTGAGAAATTTCTGAAGCCGTAATTTTATTATCAGTGAGTCAGTCATCAGGAGA	4800
Qy	4801	TCCTCTACGGCGGACGATCGTGGCCGACCTGCAGGTCGGCATCACCGGCGGCACACAGTG	4860
Db	4801	TCCTCTACGGCGGACGATCGTGGCCGACCTGCAGGTCGGCATCACCGGCGGCACACAGTG	4860
Qy	4861	CGGTTGCTGGCGCCTATATCGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCG	4920
Db	4861	CGGTTGCTGGCGCCTATATCGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCG	4920
Qy	4921	GGCTCATAGCGCTGTTTTCGGCGTGGGTATGFTGGCAGGCCCGCTGGCCGGGGGACTGT	4980
Db	4921	GGCTCATAGCGCTGTTTTCGGCGTGGGTATGFTGGCAGGCCCGCTGGCCGGGGGACTGT	4980
Qy	4981	TGGCGGCATCTCCTCTGCATGCACCATTCCTTGGCGGGCGGCTGCTCAACGGCCTCAACC	5040
Db	4981	TGGCGGCATCTCCTCTGCATGCACCATTCCTTGGCGGGCGGCTGCTCAACGGCCTCAACC	5040
Qy	5041	TACTACTGGGCTGCTTCTCTAATCAGGAGTCGCATAAGGAGAGCGTCGAGTACTTATGA	5100
Db	5041	TACTACTGGGCTGCTTCTCTAATCAGGAGTCGCATAAGGAGAGCGTCGAGTACTTATGA	5100
Qy	5101	TTGGAAGTATGGGAATGGTGATACCCGATTCCTCAGTGCTTTCAGGTCCTCATACAGT	5160
Db	5101	TTGGAAGTATGGGAATGGTGATACCCGATTCCTCAGTGCTTTCAGGTCCTCATACAGT	5160
Qy	5161	TATGCCCAACTTAAGCAACCGGAGGAGAGATTTTCATGGTAAATTTCTCGACTTTTGGT	5220
Db	5161	TATGCCCAACTTAAGCAACCGGAGGAGAGATTTTCATGGTAAATTTCTCGACTTTTGGT	5220
Qy	5221	CATCAGTAGACTCGAACTGTGAGACTATCTCGGTTATGACAGCAGAAATGTCTTCTTTGG	5280
Db	5221	CATCAGTAGACTCGAACTGTGAGACTATCTCGGTTATGACAGCAGAAATGTCTTCTTTGG	5280
Qy	5281	AGACAGTAATTAAGTCCCACCAATAAGAAAATCCCTTTGTTATCAGGAACAAACTTCTTGT	5340

5281	AGACAGTAAATGAAGTCCCAACCAATAAAGAAATCCCTGTATCAGGAACAAACCTTCCTGTG	5340
5341	TTTCAACATTTTTCGGTGCCTTGAACATATAAAATGTAGAGTGGATATGTCGGGTAGGAATG	5400
5341	TTTCAACATTTTTCGGTGCCTTGAACATATAAAATGTAGAGTGGATATGTCGGGTAGGAATG	5400
5401	GAGCGGGCAAAATGCTTACCTTCTGGACCTTCAAGAGGTATGTAGGGTTTGTAGATACTGA	5460
5401	GAGCGGGCAAAATGCTTACCTTCTGGACCTTCAAGAGGTATGTAGGGTTTGTAGATACTGA	5460
5461	TGCCAACTTCAGTGACAAACGTTGCTATTTCGTTTCAAAACCATTCGGAATCCAGAGAAATCA	5520
5461	TGCCAACTTCAGTGACAAACGTTGCTATTTCGTTTCAAAACCATTCGGAATCCAGAGAAATCA	5520
5521	AAGTTGTTTGTCTACTATTGATCCAAAGCCAGTGGCGTCTTGAAACTGCAAAATAGTGTGCT	5580
5521	AAGTTGTTTGTCTACTATTGATCCAAAGCCAGTGGCGTCTTGAAACTGCAAAATAGTGTGCT	5580
5581	CGTGTTTGTAGGTCATCTTTGTATGAATAAATCTAGTCTTTGATCTAAATAATCTTCACG	5640
5581	CGTGTTTGTAGGTCATCTTTGTATGAATAAATCTAGTCTTTGATCTAAATAATCTTCACG	5640
5641	AGCCAAAGCGGATAAATACCCAAATCTTTTAAACGTTTAAAGGACAAAGTATG	5700
5641	AGCCAAAGCGGATAAATACCCAAATCTTTTAAACGTTTAAAGGACAAAGTATG	5700
5701	TCGCTCTGTATTAAACCCCAAAATCAGCTCGTAGTCTGATCTCATCAACTGTAGGGGCAC	5760
5701	TCGCTCTGTATTAAACCCCAAAATCAGCTCGTAGTCTGATCTCATCAACTGTAGGGGCAC	5760
5761	TATCTTGTTTGTAGAGAAATTCGGGAGATGCCATATCGAGAAAAAGGTACGCTGATTTTA	5820
5761	TATCTTGTTTGTAGAGAAATTCGGGAGATGCCATATCGAGAAAAAGGTACGCTGATTTTA	5820
5821	AACGTGAAATTTATCTCAAGATCTTCGCTCGCGGTTTCGGTGATGACCGGTGAAAACCT	5880
5821	AACGTGAAATTTATCTCAAGATCTTCGCTCGCGGTTTCGGTGATGACCGGTGAAAACCT	5880
5881	CTGACACATGCAGCTCCGGAGACGCTCACAGCTGTCTGTAAAGCGGATGCCGGAGCAG	5940
5881	CTGACACATGCAGCTCCGGAGACGCTCACAGCTGTCTGTAAAGCGGATGCCGGAGCAG	5940
5941	ACAAGCCCGTCAGGCGCGCTCAGCGGTTGTTCGGCGGTTTCGGGCGCAGCCATGACCCA	6000
5941	ACAAGCCCGTCAGGCGCGCTCAGCGGTTGTTCGGCGGTTTCGGGCGCAGCCATGACCCA	6000
6001	GTCAGTAGCGATAGCGGAGTGTATCTGGCTTAATCTACCTATTCGGCATCAGAGCAGATTGA	6060
6001	GTCAGTAGCGATAGCGGAGTGTATCTGGCTTAATCTACCTATTCGGCATCAGAGCAGATTGA	6060
6061	CTGAGGTGCAACCATATGCGGTTGAAATACCCGCAAGATTCGTAAGGAGAAAAATACCGC	6120
6061	CTGAGGTGCAACCATATGCGGTTGAAATACCCGCAAGATTCGTAAGGAGAAAAATACCGC	6120
6121	ATCAGGCGCTCTTCGGTTCCTCGTCACTGACTGCTGCGCTCGGCTCGGCTCGGCTCGG	6180
6121	ATCAGGCGCTCTTCGGTTCCTCGTCACTGACTGCTGCGCTCGGCTCGGCTCGGCTCGG	6180
6181	CGAGCGGTATCAGCTCACTCAAAGCGGTAAATACCGTATCCACAGAAATCAGGGGATAAC	6240
6181	CGAGCGGTATCAGCTCACTCAAAGCGGTAAATACCGTATCCACAGAAATCAGGGGATAAC	6240
6241	GCAGAAAGAACATGTGAGCAAAAGCCAGCAAAAGCCAGCAACCGTAAAGGCCGCG	6300
6241	GCAGAAAGAACATGTGAGCAAAAGCCAGCAAAAGCCAGCAACCGTAAAGGCCGCG	6300
6301	TTGCTGGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCAAAAAATTCACAGCTCA	6360
6301	TTGCTGGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCAAAAAATTCACAGCTCA	6360
6361	AGTCAGAGGTGGCAAAACCCGACAGGACTATTAAGATACAGGCGTTTCCTCCCTTGAAGC	6420

Db 6361 AGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACACAGGGGTTTCCCCCTGGAAAGC 6420
Qy 6421 TCCCTCGTGCCTCTCTGTTTCCGACCTCGCGCTTACCAGATACCTGTCCGCTTTCTC 6480
Db 6421 TCCCTCGTGCCTCTCTGTTTCCGACCTCGCGCTTACCAGATACCTGTCCGCTTTCTC 6480
Qy 6481 CTTTCGGGAAGCTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTGGTGTAG 6540
Db 6481 CTTTCGGGAAGCTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTGGTGTAG 6540
Qy 6541 GTCGTTTCCTCAAGCTGGGCTGTGTGACGAACCCCGCTCAGCCCGACCGCTGGCC 6600
Db 6541 GTCGTTTCCTCAAGCTGGGCTGTGTGACGAACCCCGCTCAGCCCGACCGCTGGCC 6600
Qy 6601 TTATCCGGTAACCTATCGCTTGAAGTTCACAAACCCGGTAAAGACAGCTTATCGCCACTGGCA 6660
Db 6601 TTATCCGGTAACCTATCGCTTGAAGTTCACAAACCCGGTAAAGACAGCTTATCGCCACTGGCA 6660
Qy 6661 GCAGCACTGTACAGGATAGCAGAGGAGTATGTAGCGGTGTCTACAGAGTCTTG 6720
Db 6661 GCAGCACTGTACAGGATAGCAGAGGAGTATGTAGCGGTGTCTACAGAGTCTTG 6720
Qy 6721 AAGTGTGGCCTAACTACGGCTACACTAGAGGACAGTATTTGGTATCTCGCTCTGCTG 6780
Db 6721 AAGTGTGGCCTAACTACGGCTACACTAGAGGACAGTATTTGGTATCTCGCTCTGCTG 6780
Qy 6781 AAGCAGTTACTCTCGGAAAGAGTTGGTAGCTTTGATCCGCGCAACAAACACCGCT 6840
Db 6781 AAGCAGTTACTCTCGGAAAGAGTTGGTAGCTTTGATCCGCGCAACAAACACCGCT 6840
Qy 6841 GGTAGCGTGGTTTTTTTGTGTAAGCAGCAGAGTATACCGCAGAAAAAAGGATCTCAA 6900
Db 6841 GGTAGCGTGGTTTTTTTGTGTAAGCAGCAGAGTATACCGCAGAAAAAAGGATCTCAA 6900
Qy 6901 GAAGATCTTTGATCTTTTCTACGGGCTCTGACGCTCAGTGGAACGAAACCTACGTTAA 6960
Db 6901 GAAGATCTTTGATCTTTTCTACGGGCTCTGACGCTCAGTGGAACGAAACCTACGTTAA 6960
Qy 6961 GGGATTTTGGTCATGAGATTAACAAAGGATCTTCACTAGATCTTTTAAATTAATAA 7020
Db 6961 GGGATTTTGGTCATGAGATTAACAAAGGATCTTCACTAGATCTTTTAAATTAATAA 7020
Qy 7021 TGAAGTTTAAATCAATCTAAAGTATATATGAGTAACTCGTCTGACAGTTACCAATGC 7080
Db 7021 TGAAGTTTAAATCAATCTAAAGTATATATGAGTAACTCGTCTGACAGTTACCAATGC 7080
Qy 7081 TTATCAGTGAGGACCTATCTCAGGATCTCTATTTTGGTTTCACTCATATAGTTGCCTGA 7140
Db 7081 TTATCAGTGAGGACCTATCTCAGGATCTCTATTTTGGTTTCACTCATATAGTTGCCTGA 7140
Qy 7141 CTCCCGCTCGTGTAGATAACTACGATACGGAGGCTTACCATCTGGCCCGCAGTCTGCA 7200
Db 7141 CTCCCGCTCGTGTAGATAACTACGATACGGAGGCTTACCATCTGGCCCGCAGTCTGCA 7200
Qy 7201 ATGATACCGGAGACCCAGCTCAGCTGCGGCTCAGATTTATCAGCAATAAACCCAGCGCC 7260
Db 7201 ATGATACCGGAGACCCAGCTCAGCTGCGGCTCAGATTTATCAGCAATAAACCCAGCGCC 7260
Qy 7261 GGAAGGCGGAGCGCAGAGTGGTCTCTGCACTTTATCGGCTCCATCCAGTCTATTAAAT 7320
Db 7261 GGAAGGCGGAGCGCAGAGTGGTCTCTGCACTTTATCGGCTCCATCCAGTCTATTAAAT 7320
Qy 7321 TGTTCGGGAAGCTAGAGTAACTAGTTGGCAGTAAATAGTTTGGCGCAACGTTGTGTC 7380
Db 7321 TGTTCGGGAAGCTAGAGTAACTAGTTGGCAGTAAATAGTTTGGCGCAACGTTGTGTC 7380
Qy 7381 ATTGCTGAGGCTCGTGGTGTACGCTCGTCTGTTTGGTATGGCTTCAATCAGCTCCGGT 7440
Db 7381 ATTGCTGAGGCTCGTGGTGTACGCTCGTCTGTTTGGTATGGCTTCAATCAGCTCCGGT 7440
Qy 7441 TCCCAACGATCAAGCGAGTTACATGATCCCCCAATGTTGTGCAAAAAAGCGTTAGCTCC 7500
Db 7441 TCCCAACGATCAAGCGAGTTACATGATCCCCCAATGTTGTGCAAAAAAGCGTTAGCTCC 7500

Qy 7501 TTCCGTTCTCCGATCGTTGTGAGAGTAAGTTGGCGGAGTGTATATCACTCATGTTATG 7560
Db 7501 TTCCGTTCTCCGATCGTTGTGAGAGTAAGTTGGCGGAGTGTATATCACTCATGTTATG 7560
Qy 7561 GCAGCACTGCATAAATTTCTTACTGTCTATGCAATCCGTAAGATGCTTTTCTGTGACTGT 7620
Db 7561 GCAGCACTGCATAAATTTCTTACTGTCTATGCAATCCGTAAGATGCTTTTCTGTGACTGT 7620
Qy 7621 GAGTACTCAACAAGTCAATTTGAGAAATAGTGTATGCGCGCACCGAGTTGCTTTCGCCG 7680
Db 7621 GAGTACTCAACAAGTCAATTTGAGAAATAGTGTATGCGCGCACCGAGTTGCTTTCGCCG 7680
Qy 7681 GGTCAACCGGGTAATACCGCGCACATACGAGAACTTTAAAGTGTCTCATTTGGA 7740
Db 7681 GGTCAACCGGGTAATACCGCGCACATACGAGAACTTTAAAGTGTCTCATTTGGA 7740
Qy 7741 AAACGTTCTTCGGGCGGAAACTCTCAAGGATCTTACCGCTTGGAGTCCAGTTTCGATG 7800
Db 7741 AAACGTTCTTCGGGCGGAAACTCTCAAGGATCTTACCGCTTGGAGTCCAGTTTCGATG 7800
Qy 7801 TAACCCACTCGTGCACCGAACTGATCTTTCAGCATCTTTTACCTTCCAGCGCTTTCGCG 7860
Db 7801 TAACCCACTCGTGCACCGAACTGATCTTTCAGCATCTTTTACCTTCCAGCGCTTTCGCG 7860
Qy 7861 TGAGCAAAAACAGGAAGGCAAAATCGGCAAAAAGGGAATAAGGCGCACACGGAATCT 7920
Db 7861 TGAGCAAAAACAGGAAGGCAAAATCGGCAAAAAGGGAATAAGGCGCACACGGAATCT 7920
Qy 7921 TGAATACTCATACTCTTCTTTCATATTTATGAGCATTTATCAGGATTTATGTTCTC 7980
Db 7921 TGAATACTCATACTCTTCTTTCATATTTATGAGCATTTATCAGGATTTATGTTCTC 7980
Qy 7981 ATGAGCGGATACATATTTGAATGATTTAGAAAAATTAACAAATAGGGGTTTCGCGCAC 8040
Db 7981 ATGAGCGGATACATATTTGAATGATTTAGAAAAATTAACAAATAGGGGTTTCGCGCAC 8040
Qy 8041 TTTTCCCGGAAAAAGTGCACCTGACGCTTAAGAAAAACCATTTATCATGACATTAACCTAT 8100
Db 8041 TTTTCCCGGAAAAAGTGCACCTGACGCTTAAGAAAAACCATTTATCATGACATTAACCTAT 8100
Qy 8101 AAAAAATAGGCTATCAGAGGCGCTTCTCTTCAAGATTAATCTCATGTTTGCACG 8160
Db 8101 AAAAAATAGGCTATCAGAGGCGCTTCTCTTCAAGATTAATCTCATGTTTGCACG 8160
Qy 8161 TTATCATCATAGCTGACTCATGTTTGTGTAATAGAGTGTGTAATAGCGGAGTCCGGAACCTG 8220
Db 8161 TTATCATCATAGCTGACTCATGTTTGTGTAATAGAGTGTGTAATAGCGGAGTCCGGAACCTG 8220
Qy 8221 AAAAAATACAGTTATTTTCG 8241
Db 8221 AAAAAATACAGTTATTTTCG 8241

RESULT 2

E16036 8598 bp DNA linear PAT 28-JUL-1999
LOCUS cdna encoding improved Pichia elafin.
DEFINITION
ACCESSION E16036
VERSION E16036.1 GI:5710719
KEYWORDS JP 1998127292-A/1.
SOURCE Pichia pastoris
ORGANISM Pichia pastoris
REFERENCE
Rukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Pichia.
1 (bases 1 to 8598)
AUTHORS Taniyama, M., Yamamoto, T., Okawa, N. and Zushi, M.
TITLE ELAFIN-EXPRESSION VECTOR AND PRODUCTION OF ELAFINS BY UTILIZING
THE SAME
JOURNAL Parent: JP 1998127292-A 1 19-MAY-1998;
TSUMURA & CO
COMMENT OS Pichia pastoris
PN JP 1998127292-A/1

PD 19-MAY-1998
 PF 31-OCT-1996 JP 1996304233
 PI TANIYAMA MITSUE, YAMAMOTO TAKASHI, OKAWA NORIYUKI, PI ZUSHI
 MAKOTO
 PC C12N15/09, C07H21/04, C07K14/39, C12N1/19, C12P21/02, (C12N1/19, PC
 C12R1/84),
 PC (C12P21/02, C12R1/84);
 CC strandedness: Double;
 CC topology: Linear;
 CC hypochromic: No;
 CC anti-sense: No;
 FH Key Location/Qualifiers

FT source 1..8598
 /organism='Pichia pastoris'
 /clone='pPIC9/BLP25L'
 FT CDS 949..1377
 /product='improved Pichia elafin' FT
 sig_peptide 949..1203
 FT mat_peptide 1204..1374
 /product='improved Pichia elafin' FT
 misc_feature replace(1276..1278,'atg')
 FT /note='native Pichia elafin'

FEATURES

source

1..8598
 /organism='Pichia pastoris'
 /mol_type='genomic DNA'
 /db_xref='taxon:4922'

ORIGIN

Query Match 91.7%; Score 7555.6; DB 6; Length 8598;

Best Local Similarity 94.0%; Pred. No. 0;

Matches 8098; Conservative 0; Mismatches 129; Indels 385; Gaps 8;

QY 1 AGATCTACATCCAAAGACGAAAGGTTGAATGAACCTTTTGCCATCCGACATCCACAG 60
 DB 1 AGATCTACATCCAAAGACGAAAGGTTGAATGAACCTTTTGCCATCCGACATCCACAG 60
 QY 61 GTCCATTCTCACATAAGTCCCAAGCAACAGAGGGGATACACTAGCAGGACCGT 120
 DB 61 GTCCATTCTCACATAAGTCCCAAGCAACAGAGGGGATACACTAGCAGGACCGT 120
 QY 121 TGCAAAACGAGACCTCCACTCTCTCTCTCTCAACACCCACTTTTGCCATCGAAAACC 180
 DB 121 TGCAAAACGAGACCTCCACTCTCTCTCTCTCAACACCCACTTTTGCCATCGAAAACC 180
 QY 181 AGCCAGTATTGGGCTTGATGGAGCTCGCTCAATTCCTTCTATTAGGCTACTA 240
 DB 181 AGCCAGTATTGGGCTTGATGGAGCTCGCTCAATTCCTTCTATTAGGCTACTA 240
 QY 241 ACACATGACTTTATTAGCTGTCTATCTCTGCGCCCTGCGAGGTTCAATTTGTTA 300
 DB 241 ACACATGACTTTATTAGCTGTCTATCTCTGCGCCCTGCGAGGTTCAATTTGTTA 300
 QY 301 TTTCGAATGCAACAGCTCCGCAATTACACCCGAACATCACTCCAGATGAGGCTTTCTG 360
 DB 301 TTTCGAATGCAACAGCTCCGCAATTACACCCGAACATCACTCCAGATGAGGCTTTCTG 360
 QY 361 AGTGTGGGTCAAATAGATTTCATGTTCCCAAAATGGGCCAAACACTGACAGTTAAAGCT 420
 DB 361 AGTGTGGGTCAAATAGATTTCATGTTCCCAAAATGGGCCAAACACTGACAGTTAAAGCT 420
 QY 421 GTCCTGGAACTAATATGACAAAAGCGTGATCTCATCCAGATGAACCTAAAGTTGGTTCG 480
 DB 421 GTCCTGGAACTAATATGACAAAAGCGTGATCTCATCCAGATGAACCTAAAGTTGGTTCG 480
 QY 481 TTGAATGCTAACGCCAGTTGGTTCGCAAAAGAAACTTCGAAAAGTCCGCATACGTTTGT 540
 DB 481 TTGAATGCTAACGCCAGTTGGTTCGCAAAAGAAACTTCGAAAAGTCCGCATACGTTTGT 540
 QY 541 CTTGTTTGGTATTGATTCACCAATGCTCAAAAATAAATCTCAATTAATGCTTAGCGAGTCT 600
 DB 541 CTTGTTTGGTATTGATTCACCAATGCTCAAAAATAAATCTCAATTAATGCTTAGCGAGTCT 600

QY 601 CTCATCGCTTCTGAAACCCCGGTGCACCTGTGCGGAAACGCAAAATGGGAAACACCCGCT 660
 DB 601 CTCATCGCTTCTGAAACCCCGGTGCACCTGTGCGGAAACGCAAAATGGGAAACACCCGCT 660
 QY 661 TTTTGGATGATTATGCAATGTCTCACAATTGTATGCTTCCAGATTCTTGGTGGAAATACT 720
 DB 661 TTTTGGATGATTATGCAATGTCTCACAATTGTATGCTTCCAGATTCTTGGTGGAAATACT 720
 QY 721 GCTGATAGCCTAAAGCTTCATGATCAAAATTTAACTGTCTAAACCCCTACTTGACAGCAAT 780
 DB 721 GCTGATAGCCTAAAGCTTCATGATCAAAATTTAACTGTCTAAACCCCTACTTGACAGCAAT 780
 QY 781 ATATAACAGAAAGGAGCTGCCCTGTCTTAAACCTTTTATCATCATATTAGCTT 840
 DB 781 ATATAACAGAAAGGAGCTGCCCTGTCTTAAACCTTTTATCATCATATTAGCTT 840
 QY 841 ACTTTTCATAATTGCGACTGGTTCCAATTTGACAAGCTTTTGAATTTTAAAGCTTTTAAAGA 900
 DB 841 ACTTTTCATAATTGCGACTGGTTCCAATTTGACAAGCTTTTGAATTTTAAAGCTTTTAAAGA 900
 QY 901 CAACTTGAGAAGATCAAAACCAACTTAATTTCGAAGGATCCAAACGATGAGATTTCCT 960
 DB 901 CAACTTGAGAAGATCAAAACCAACTTAATTTCGAAGGATCCAAACGATGAGATTTCCT 960
 QY 961 TCAATTTTACTGCAAGTTTATTCGAGCATCTCCGCAATTCCTCCAGTCCAGTCAACACT 1020
 DB 961 TCAATTTTACTGCAAGTTTATTCGAGCATCTCCGCAATTCCTCCAGTCCAGTCAACACT 1020
 QY 1021 ACAACAGAAAGATGAAACCGCACAAATCCCGCTGAAAGCTGTCATCGGTTACTCAGATTTA 1080
 DB 1021 ACAACAGAAAGATGAAACCGCACAAATCCCGCTGAAAGCTGTCATCGGTTACTCAGATTTA 1080
 QY 1081 GAAGGGGATTTGCGATGTTGCTGTTTGGCCATTTTCCAAACAGCACAAATAACGGGTTATTG 1140
 DB 1081 GAAGGGGATTTGCGATGTTGCTGTTTGGCCATTTTCCAAACAGCACAAATAACGGGTTATTG 1140
 QY 1141 TTTTAAATAACTACTACTATTGCCAGCATTCCTGCTCTAAAGAAAGGGGTATCTCTCGAGAAA 1200
 DB 1141 TTTTAAATAACTACTACTATTGCCAGCATTCCTGCTCTAAAGAAAGGGGTATCTCTCGAGAAA 1200
 QY 1201 AGAG--AGGCTGAAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 1257
 DB 1201 AGAGCTCAAGAACCACTTAAGGGTCCGGTTCGACCAACCCGGCTCTTGCCCGATTATC 1260
 QY 1258 ATGGCCCCCGGAGAGGATTAATTTGGCTTCCCGGTGTCACGCCAGCCAGCCAGCCAGCCAG 1317
 DB 1261 CTGATCCCGCTGCGCTTTGCTGAACCCGCGAACCGTTGCTGAAGACACTGACTGCGCG 1320
 QY 1318 GAG----- 1320
 DB 1321 GGTATCAAAAATGCTCGAAAGTTCTTTCGGTATGGCATGCTTCGTTCCGAGTAGTGA 1380
 QY 1321 ----- 1320
 DB 1381 AGCTCTCGAGGCCAGCTTGCGTGCAGTGCAGGAAGAACCCATCTCGCTTAATGAGTAA 1440
 QY 1321 ----- 1320
 DB 1441 TATCTGATAAGAACGCTTTTATTTAGCTTTTAACTACATAGACGAATAATAATACGA 1500
 QY 1321 ----- 1320
 DB 1501 TCGGATACATTAGCATATTAGCTCTTTTACCTCTCAATTTGCTCAAGATTAGTATCGT 1560
 QY 1321 ----- 1320
 DB 1561 GACCAATTTTAGATCATTTTCGTCAGAACAGCCACATTTTCGCGTAAATTTTAAAGTTCAA 1620
 QY 1321 -----AGAGGTGCTGTTTGTGATGACAGTGTCCGGGAT 1354
 DB 1621 TTATTGTATGCTGACACTATATTAGCGTCAACTGTTTAAATAGGAAGTAGTGGATAAT 1680

QY 1355 TCCGTTGGTCTTCCACCCCATGCGATCGAGAACACTCAAGA-----AGAGAATGTC 1409
Db 1681 AACGGTTACAGTGTGTTACAGATGCTTATTTCAAAATCTAGATGTTTAAACATTTGG 1740
QY 1410 CTTCTPACTATGGGGT-----AGAA 1430
Db 1741 ATACTAACGACATAGTTCTCCTGGCATCTGCCACCGTTGATATAGTAAACTACAAGAA 1800
QY 1431 TTCCTAGGGCGCGCGAAATAATTCGCTTAGACATGCTGTTCTCAGTTCAAGTTG 1490
Db 1801 TTCCTAGGGCGCGCGAAATAATTCGCTTAGACATGCTGTTCTCAGTTCAAGTTG 1860
QY 1491 GGCATCTACGAGAAGACCGGCTTCTGTAGATTCTAATCAAGAGGATGTCAGAAFGCCATT 1550
Db 1861 GGCATCTACGAGAAGACCGGCTTCTGTAGATTCTAATCAAGAGGATGTCAGAAFGCCATT 1920
QY 1551 TGCCTGAGAGATGAGGGTTCATTTTTCATCTTTTATTTGTAACCTATATAGTATAG 1610
Db 1921 TGCCTGAGAGATGAGGGTTCATTTTTCATCTTTTATTTGTAACCTATATAGTATAG 1980
QY 1611 GATTTTTTTTGTCAATTTGTTTCTCTCGTAGAGCTTGCTCTGATCAGCCTATCTCGC 1670
Db 1981 GATTTTTTTTGTCAATTTGTTTCTCTCGTAGAGCTTGCTCTGATCAGCCTATCTCGC 2040
QY 1671 AGCTGATGAATATCTTGTGTAGGGTTTGGGAAAAATCATTCGAGTTTGATGTTTCTT 1730
Db 2041 AGCTGATGAATATCTTGTGTAGGGTTTGGGAAAAATCATTCGAGTTTGATGTTTCTT 2100
QY 1731 GGTATTTCCCACTCTCTTCAGAGTACAGAGAATTAAGTGAGAAAGTTGTTTGGCAAGC 1790
Db 2101 GGTATTTCCCACTCTCTTCAGAGTACAGAGAATTAAGTGAGAAAGTTGTTTGGCAAGC 2160
QY 1791 TTATCGATPAAGCTTTPAATCGGTAAGTTTATCACAGTTAAATTTGTAACGCACTCAGGCAC 1850
Db 2161 TTATCGATPAAGCTTTPAATCGGTAAGTTTATCACAGTTAAATTTGTAACGCACTCAGGCAC 2220
QY 1851 CGTGATGAATCTPAACATCGGCTCATCGTCATCTCGGACCGTCCAGCTGATGCTG 1910
Db 2221 CGTGATGAATCTPAACATCGGCTCATCGTCATCTCGGACCGTCCAGCTGATGCTG 2280
QY 1911 TAGGCATAGGCTTGTATTGCGGTACTGCGGGGCTCTTGGGGGATATGTCATTCGG 1970
Db 2281 TAGGCATAGGCTTGTATTGCGGTACTGCGGGGCTCTTGGGGGATATGTCATTCGG 2340
QY 1971 ACAGATCGCAGTCACTATGCGGTGCTGTAGCGCTATATGCTTGATGCAATTTCTAT 2030
Db 2341 ACAGATCGCAGTCACTATGCGGTGCTGTAGCGCTATATGCTTGATGCAATTTCTAT 2400
QY 2031 GCGCACCGGTTCTCGGAGCACTGTCCGACCGCTTGGCGCGCGCCAGCTCGCTCGCTT 2090
Db 2401 GCGCACCGGTTCTCGGAGCACTGTCCGACCGCTTGGCGCGCGCCAGCTCGCTCGCTT 2460
QY 2091 CGCTACTTGGAGCCACTATCGACTACGCGATCATGGCGACCAACCCGCTCTGTGGATCT 2150
Db 2461 CGCTACTTGGAGCCACTATCGACTACGCGATCATGGCGACCAACCCGCTCTGTGGATCT 2520
QY 2151 ATCGAATCTAATGTAAGTAAATCTCTAATAATTAATTAAGTCCAGTTTCTCCATA 2210
Db 2521 ATCGAATCTAATGTAAGTAAATCTCTAATAATTAATTAAGTCCAGTTTCTCCATA 2580
QY 2211 CGAACCTTAAACAGATGCGGTGAGCATCTAGACCTTCAACAGCAGCGAGATCCATCACT 2270
Db 2581 CGAACCTTAAACAGATGCGGTGAGCATCTAGACCTTCAACAGCAGCGAGATCCATCACT 2640
QY 2271 GCTTGGCCAAATATGTTTACGCTCCCTCAGGATGATGCTTGTGAGTGAATGAACTTCTGG 2330
Db 2641 GCTTGGCCAAATATGTTTACGCTCCCTCAGGATGATGCTTGTGAGTGAATGAACTTCTGG 2700
QY 2331 AAGGTTGAGTGTAACTCCGCTGATTTGACGGGATATCCGTAGCTTGGCAAGTGTGG 2390
Db 2701 AAGGTTGAGTGTAACTCCGCTGATTTGACGGGATATCCGTAGCTTGGCAAGTGTGG 2760
QY 2391 TTGGTACCGGAGGATTAATCTCCACAACTCTCTGAGAGTAGGACCAACAAACACAGAT 2450

Db 2761 TTGGTACCGGAGGATTAATCTCCAACTCTCTGGAGAGTAGGCACCAACAAACACAGAT 2820
QY 2451 CAGCGGTGTTGATCTGATCAATAGAGAGCATCTTCGATTTGCAAGGATCAAGTGT 2510
Db 2821 CAGCGGTGTTGATCTGATCAATAGAGAGCATCTTCGATTTGCAAGGATCAAGTGT 2880
QY 2511 TCAGGAGCCGTAAGTTGAGCATTTCCAAAGCCTCTCTAGTGTGCAACCGATAGGGTT 2570
Db 2881 TCAGGAGCCGTAAGTTGAGCATTTCCAAAGCCTCTCTAGTGTGCAACCGATAGGGTT 2940
QY 2571 GTAGAGTGTCAATACACTTGGGTACAAATTTCAAACCTTTGGCAACTGCAAGCTGGTTG 2630
Db 2941 GTAGAGTGTCAATACACTTGGGTACAAATTTCAAACCTTTGGCAACTGCAAGCTGGTTG 3000
QY 2631 TGAACAGCATCTTCAATCTGCAAGCTCTTGTCTGTCATATCGACACCAACAGAATC 2690
Db 3001 TGAACAGCATCTTCAATCTGCAAGCTCTTGTCTGTCATATCGACAGCCAAAGATC 3060
QY 2691 ACCTGGGAATCAATACCATGTTGAGCTTGAGACAGAAAGTCTGAGGCAACGAAATCTGGA 2750
Db 3061 ACCTGGGAATCAATACCATGTTGAGCTTGAGCTTGAGGCAACGAAATCTGGA 3119
QY 2751 TCAGCGTATTTATCAGCAATTAAGTAACTTCAGAGGCCCGCAGCGGATGTCATACTA 2810
Db 3120 TCAGCGTATTTATCAGCAATTAAGTAACTTCAGAGGCCCGCAGCGGATGTCATACTA 3179
QY 2811 CACAGGCGTGTGTGTCATTTTGAACCATCATCTTTGGCAGCAGTAAACGAATGTTTCT 2870
Db 3180 CACAGGCGTGTGTGTCATTTTGAACCATCATCTTTGGCAGCAGTAAACGAATGTTTCT 3239
QY 2871 GGAACCAATATTTGTCACACTTAGGAACAGTTTCTGTTCCGTAAAGCCATAGCAGTACT 2930
Db 3240 GGAACCAATATTTGTCACACTTAGGAACAGTTTCTGTTCCGTAAAGCCATAGCAGTACT 3299
QY 2931 GCGTGGCGCTCTCTCTAGCAGTACACTTTAGCACCAACCTTGTGGGCAACGATAGTG 2990
Db 3300 GCGTGGCGCTCTCTCTAGCAGTACACTTTAGCACCAACCTTGTGGGCAACGATAGTG 3359
QY 2991 ACTTCTGGGTAAAGGTACATCTCTTTAGGTGAGATGCAAAAACAATTTCTTTGCAA 3050
Db 3360 ACTTCTGGGTAAAGGTACATCTCTTTAGGTGAGATGCAAAAACAATTTCTTTGCAA 3419
QY 3051 CAGCAACTTTGGCAGGACACCCAGCATCAGGAGTGAAGGCAAGATTCGGTTCCA 3110
Db 3420 CAGCAACTTTGGCAGGACACCCAGCATCAGGAGTGAAGGCAAGATTCGGTTCCA 3479
QY 3111 CCAGGAATATAGAGGCCAACTTTCTCAATAGGTCTTTGCAAAAACGAGAGCAGACTACCCA 3170
Db 3480 CCAGGAATATAGAGGCCAACTTTCTCAATAGGTCTTTGCAAAAACGAGAGCAGACTACCCA 3539
QY 3171 GGCAGATCTCAACTTTGCAACGCTCTCGTGTAGTGTGAGCTTCATGGAATTTCTGAGCTTA 3230
Db 3540 GGCAGATCTCAACTTTGCAACGCTCTCGTGTAGTGTGAGCTTCATGGAATTTCTGAGCTTA 3599
QY 3231 TCTATAGAGATCAATGGCTCTCTTAAAGCTTATCTTGGCAATTTGCAATAGTTCTCTGG 3290
Db 3600 TCTATAGAGATCAATGGCTCTCTTAAAGCTTATCTTGGCAATTTGCAATAGTTCTCTGG 3659
QY 3291 AAAGGAGCTTAAACAGAGTGTCTTCAAGGCACTCCATCAAACTTGGCAGTGTGTTCT 3350
Db 3660 AAAGGAGCTTAAACAGAGTGTCTTCAAGGCACTCCATCAAACTTGGCAGTGTGTTCT 3719
QY 3351 AAAGGAGCTTGTACCACTTTGACCAACATTTGCGACAAATTTGTTGACTAATTCCTA 3410
Db 3720 AAAGGAGCTTGTACCACTTTGACCAACATTTGCGACAAATTTGTTGACTAATTCCTA 3779
QY 3411 ATCTGTTCCGTTTCTTGATAGGACGCAAGGCACTCTTCAATTTCTTGTGAGGAGGCC 3470
Db 3780 ATCTGTTCCGTTTCTTGATAGGACGCAAGGCACTCTTCAATTTCTTGTGAGGAGGCC 3839
QY 3471 TTAGAAAGCTCAATTTTGCACAAATTTCAATACGACCTTCAGAGGAGGACTCTTTAGTTTG 3530

Db 3840 TTAGAAACGTCATTTTGGCAAAATCAATAGACCTTTCAGAAAGGACTCTTTTAGGTTTG 3899
Qy 3531 GATTCCTCTTTAGGTTGTTCTTGGTGTATCCCTGGCTTGGCATCTCTTTCTTCTAGTG 3590
Db 3900 GATTCCTCTTTAGGTTGTTCTTGGTGTATCCCTGGCTTGGCATCTCTTTCTTCTAGTG 3959
Qy 3591 ACCTTTAGGAGCTTCATATCAGAGTTTCTCTCACCTCTGTCGAACGTACACCGTACTTG 3650
Db 3960 ACCTTTAGGAGCTTCATATCAGAGTTTCTCTCACCTCTGTCGAACGTACACCGTACTTG 4019
Qy 3651 GCACATCTAAGTAAATGCAAAATAAATAAGTCAGACATTTCCAGGCTATATCTTCTCTTG 3710
Db 4020 GCACATCTAAGTAAATGCAAAATAAATAAGTCAGACATTTCCAGGCTATATCTTCTCTTG 4079
Qy 3711 GATTTAGCTTCTGCAAGTTCAATAGCTTCTCTCTAATTTTAGGTTTC-AACAAACTTC 3769
Db 4080 GATTTAGCTTCTGCAAGTTCAATAGCTTCTCTCTAATTTTAGGTTTCACAAACAAACTTC 4139
Qy 3770 GTCGTCAAAATAACCGTTTGGTATAGAACCTTCTGGAGCATTTGCTCTTAGATCCCAAA 3829
Db 4140 GTCGTCAAAATAACCGTTTGGTATAGAACCTTCTGGAGCATTTGCTCTTAGATCCCAAA 4199
Qy 3830 GGTGGCTTCCATGGCTCTAAGACCTTTTGATTTGGCCAAACACAGAAAGTGGTTTCCAAAGT 3889
Db 4200 GGT-GCTTCCATGGCTCTAAGACCTTTTGATTTGGCCAAACACAGAAAGTGGTTTCCAAAGT 4258
Qy 3890 ACAGAAACCAACCTGTTTCAACCAACAAATTTCAAGCAGTCTCCATCAAAATCCAA 3949
Db 4259 ACAGAAACCAACCTGTTTCAACCAACAAATTTCAAGCAGTCTCCATCAAAATCCAA 4318
Qy 3950 TTCCATATCCAGCACTTTTGGATTTGCTCCAGATTTAGTGTAGCACTTTTATACCAAAACCGTG 4009
Db 4319 TTCCATATCCAGCACTTTTGGATTTGCTCCAGATTTAGTGTAGCACTTTTATACCAAAACCGTG 4378
Qy 4010 ACAGAGATTTGGTAGCTCAGTTTGTGCTTATAGCTTCCGGATAGACTTTTGGTA 4069
Db 4379 ACAGAGATTTGGTAGCTCAGTTTGTGCTTATAGCTTCCGGATAGACTTTTGGTA 4438
Qy 4070 CGAGTACACAGGCGCCAAAGAGTAAATAGAGAGTTCAGCCACCAAGTAGTGAATAGAC 4129
Db 4439 CGAGTACACAGGCGCCAAAGAGTAAATAGAGAGTTCAGCCACCAAGTAGTGAATAGAC 4498
Qy 4130 ATCCGGGCGGTGAGTCAAGAGCGCAACAAATTTCACTGACAGGGAACCTTTTGGAC 4189
Db 4499 ATCCGGGCGGTGAGTCAAGAGCGCAACAAATTTCACTGACAGGGAACCTTTTGGAC 4558
Qy 4190 ATCTTCAGAAAGTTCGATTCAGTAGTCAATTCGCGAGCATCAATAATGGGATTATACC 4249
Db 4559 ATCTTCAGAAAGTTCGATTCAGTAGTCAATTCGCGAGCATCAATAATGGGATTATACC 4618
Qy 4250 AGAAGCAACAGTGGAGTCACTATACCACTTTTGGGCTCTCAGAAAAAGCATAAACAGT 4309
Db 4619 AGAAGCAACAGTGGAGTCACTATACCACTTTTGGGCTCTCAGAAAAAGCATAAACAGT 4678
Qy 4310 TCTACTACCGCATTTAGTGAACCTTTCAATTCGCGAGTGGAGAAAGGACAGC 4369
Db 4679 TCTACTACCGCATTTAGTGAACCTTTCAATTCGCGAGTGGAGAAAGGACAGC 4738
Qy 4370 GATATAGCATTTAGCGGCGAGGATGCAACCTTTATCAACAGGCTCTTATAGATAAACCTT 4429
Db 4739 GATATAGCATTTAGCGGCGAGGATGCAACCTTTATCAACAGGCTCTTATAGATAAACCTT 4798
Qy 4430 AGCGCTGGATCATCTTTTGGACAACTCTTTTGGCCAAATCTAGGTCCCAAAATCACTTC 4489
Db 4799 AGCGCTGGATCATCTTTTGGACAACTCTTTTGGCCAAATCTAGGTCCCAAAATCACTTC 4858
Qy 4490 ATTGTATACATTTATTTGTAACCTTGGACAACTTTGCGATCAGCTCTCAAAATGGTCTTC 4549
Db 4859 ATTGTATACATTTATTTGTAACCTTGGACAACTTTGCGATCAGCTCTCAAAATGGTCTTC 4918
Qy 4550 TGTAACCGATGACTCAACTTTGACAACTTTAACTTTGAAGCTCAGTCGATGAGTGAACCTTGAT 4609
Db 4919 TGTAACCGATGACTCAACTTTGACAACTTTAACTTTGAAGCTCAGTCGATGAGTGAACCTTGAT 4978

Qy 4610 CAGTTGTGCACTGGTCTCAGCAGCATAGGAAACACCGCTTTTCTACCAAACTCAAGGA 4669
Db 4979 CAGTTGTGCACTGGTCTCAGCAGCATAGGAAACACCGCTTTTCTACCAAACTCAAGGA 5038
Qy 4670 ATTATCAAACTCTGCAACACCTTTGGGTATGCAAGGTAGCAAGGAAATGTCACTTTGAAGT 4729
Db 5039 ATTATCAAACTCTGCAACACCTTTGGGTATGCAAGGTAGCAAGGAAATGTCACTTTGAAGT 5098
Qy 4730 CGGACAGTGTAGTGTCTTGAGAAATTTCTGAACCGTATTTTATATCATCAGTCAGTCAG 4789
Db 5099 CGGACAGTGTAGTGTCTTGAGAAATTTCTGAACCGTATTTTATATCATCAGTCAGTCAG 5158
Qy 4790 TCATCAGAGATCTCTACGCGGACGATCTGTCGCGACCTGCGAGTTCGGCATCACCGG 4849
Db 5159 TCATCAGAGATCTCTACGCGGACGATCTGTCGCGACCTGCGAGTTCGGCATCACCGG 5206
Qy 4850 CGCCAAGTGTGGTGTCTGCGGCTATATCGCGACATCAACGATGGGAGAGATCGGGC 4909
Db 5207 CGCCAAGTGTGGTGTCTGCGGCTATATCGCGACATCAACGATGGGAGAGATCGGGC 5266
Qy 4910 TCGCCACTTCGGGCTCATGAGCGCTTTTTCGGGTGGGTATGTTGGCAGGCCCTGTCG 4969
Db 5267 TCGCCACTTCGGGCTCATGAGCGCTTTTTCGGGTGGGTATGTTGGCAGGCCCTGTCG 5326
Qy 4970 CGGGGGACTGTTGGGCGCCATCTCTTTCATGCAACCATTCCTTTCGGGCGGGTGTCAA 5029
Db 5327 CGGGGGACTGTTGGGCGCCATCTCTTTCATGCAACCATTCCTTTCGGGCGGGTGTCAA 5386
Qy 5030 CGGCCCTCAACCTTACTTCTGCGGCTGCTTCTTAATCGAGGATCGCATAGGAGAGAGCTCG 5089
Db 5387 CGGCCCTCAACCTTACTTCTGCGGCTGCTTCTTAATCGAGGATCGCATAGGAGAGAGCTCG 5446
Qy 5090 AGTATCTATGATTGGAAGTATGGGAATGGTATACCCGCTTTCTTTCAGTGTCTGAAGTC 5149
Db 5447 AGTATCTATGATTGGAAGTATGGGAATGGTATACCCGCTTTCTTTCAGTGTCTGAAGTC 5506
Qy 5150 TCCTATCAGATTAATGCCCAACTAAAGCAACCGAGGAGAGATTTTCATGTTAAATTTCTC 5209
Db 5507 TCCTATCAGATTAATGCCCAACTAAAGCAACCGAGGAGAGATTTTCATGTTAAATTTCTC 5566
Qy 5210 TGACTTTTGGTCTCATCAGTAGACTCGAACTGTGAGACTATCTCGGTATGACAGCAAAAT 5269
Db 5567 TGACTTTTGGTCTCATCAGTAGACTCGAACTGTGAGACTATCTCGGTATGACAGCAAAAT 5626
Qy 5270 GTCTTCTTGGAGACAGTAAATGAAAGTCCCACTAAAGAAATCCCTTGTATCAGGAAAC 5329
Db 5627 GTCTTCTTGGAGACAGTAAATGAAAGTCCCACTAAAGAAATCCCTTGTATCAGGAAAC 5686
Qy 5330 AAACCTTGTGTTTCCAACTTTTTCGGTGGCTTGAATATAAACTGAGTGGATATGTC 5389
Db 5687 AAACCTTGTGTTTCCAACTTTTTCGGTGGCTTGAATATAAACTGAGTGGATATGTC 5746
Qy 5390 GGGTAGGAATGAGCGGCAAAATGCTTACCTTCTGACCTTCAAGAGGTATGTTAGGGTTT 5449
Db 5747 GGGTAGGAATGAGCGGCAAAATGCTTACCTTCTGACCTTCAAGAGGTATGTTAGGGTTT 5806
Qy 5450 GTAGATCTGATGCCAACTTTCAGTACCAACGTTGCTATTTCTGTTCAAACTTCCGAATC 5509
Db 5807 GTAGATCTGATGCCAACTTTCAGTACCAACGTTGCTATTTCTGTTCAAACTTCCGAATC 5866
Qy 5510 CAGAGAAATCAAAAGTGTGTTTGTCTACTATTTGATCCCAAGCCAGTGGCTTCTGAAACTGAC 5569
Db 5867 CAGAGAAATCAAAAGTGTGTTTGTCTACTATTTGATCCCAAGCCAGTGGCTTCTGAAACTGAC 5926
Qy 5570 AATAGTGTCTCGTGTGTTTGGTGTCTTCTTGTATGAATTAATCTAGTCTTGTGATCTAAA 5629
Db 5927 AATAGTGTCTCGTGTGTTTGGTGTCTTCTTGTATGAATTAATCTAGTCTTGTGATCTAAA 5986
Qy 5630 TAATCTTTCAGGAGCAAGGCGATATAACCCAAATCTTAAACTCTTTTAAACCGTTAAAA 5689
Db 5987 TAATCTTTCAGGAGCAAGGCGATATAACCCAAATCTTAAACTCTTTTAAACCGTTAAAA 6046

QY 5690 GGACAGATATGTCCTCTGTATTAAACCCCAATCAGCTCGTAGTCTGATCCTCATCAAC 5749
DB 6047 GGACAGATATGTCCTCTGTATTAAACCCCAATCAGCTCGTAGTCTGATCCTCATCAAC 6106
QY 5750 TTGAGGGCAGCTATCTCTTTAGAGAAATTTGGGAGATCGGATATCGAGAAAAGGTA 5809
DB 6107 TTGAGGGCAGCTATCTCTTTAGAGAAATTTGGGAGATCGGATATCGAGAAAAGGTA 6166
QY 5810 CGCTGATTTAAACGCTGAAATTTATCTCAAGATCTCTGCTCGCGGTTTCGGTGATGAC 5869
DB 6167 CGCTGATTTAAACGCTGAAATTTATCTCAAGATCTCTGCTCGCGGTTTCGGTGATGAC 6226
QY 5870 GGTGAARACCTTGACACATGACAGTCCGAGAGCGGTACAGCTTGTCTGTAAAGCGAT 5929
DB 6227 GGTGAARACCTTGACACATGACAGTCCGAGAGCGGTACAGCTTGTCTGTAAAGCGAT 6286
QY 5930 GCCGGGAGCAGACAGCCCGTCAGCGGCGGTGAGCGGCGGTGCGGGGCGCA 5989
DB 6287 GCCGGGAGCAGACAGCCCGTCAGCGGCGGTGAGCGGCGGTGCGGGGCGCA 6346
QY 5990 GCATGACCCAGTCACTGAGGATAGCGGAGTGTATGCTGTTAACTATGCGGCATCAG 6049
DB 6347 GCATGACCCAGTCACTGAGGATAGCGGAGTGTATGCTGTTAACTATGCGGCATCAG 6406
QY 6050 AGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAATACCGCACAGATGCGTAAGGA 6109
DB 6407 AGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAATACCGCACAGATGCGTAAGGA 6466
QY 6110 GAAATACCGCATCAGGCGCTCTTCGCTTCTCGCTCACTGCTGCGTGCCTGCGTGC 6169
DB 6467 GAAATACCGCATCAGGCGCTCTTCGCTTCTCGCTCACTGCTGCGTGCCTGCGTGC 6526
QY 6170 TTCGCTCGGCGAGCGGTATCAGTCACTCAAGCGGTTAATACGTTATTCACAGAT 6229
DB 6527 TTCGCTCGGCGAGCGGTATCAGTCACTCAAGCGGTTAATACGTTATTCACAGAT 6586
QY 6230 CAGGGGATAACGAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGAACCGTA 6289
DB 6587 CAGGGGATAACGAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGAACCGTA 6646
QY 6290 AAAAGGCGCGTGTCTGCGGTTTTTCATAGGCTCCGCGCCCTGAGCAGGATCACAAA 6349
DB 6647 AAAAGGCGCGTGTCTGCGGTTTTTCATAGGCTCCGCGCCCTGAGCAGGATCACAAA 6706
QY 6350 ATCGACGCTCAAGTCAGAGGTGGGAAACCGACAGACTATAAAGATACCGAGCGTTTC 6409
DB 6707 ATCGACGCTCAAGTCAGAGGTGGGAAACCGACAGACTATAAAGATACCGAGCGTTTC 6766
QY 6410 CCCCTGGAAGTCCCTCGTGGCTTCTCTGTTCCGACCCCTCCGCTTAACCGGATACCTGT 6469
DB 6767 CCCCTGGAAGTCCCTCGTGGCTTCTCTGTTCCGACCCCTCCGCTTAACCGGATACCTGT 6826
QY 6470 CCGCTTTCTCCCTTCGGGAGCGTGGCGCTTCTCATAGCTCAGCTGAGGTATCTCA 6529
DB 6827 CCGCTTTCTCCCTTCGGGAGCGTGGCGCTTCTCATAGCTCAGCTGAGGTATCTCA 6886
QY 6530 GTTCGGTGTAGTCTGCTCCAGCTGGGCTGTGTGCAAGAACCCCGCTTCAGGCGG 6589
DB 6887 GTTCGGTGTAGTCTGCTCCAGCTGGGCTGTGTGCAAGAACCCCGCTTCAGGCGG 6946
QY 6590 ACCGCTGGCCTTATCCGGTAACTATCGTCTTGAGTCCAAACCCGGTAAAGACAGCTTAT 6649
DB 6947 ACCGCTGGCCTTATCCGGTAACTATCGTCTTGAGTCCAAACCCGGTAAAGACAGCTTAT 7006
QY 6650 CGCCACTGGCAGCAGCAGTGTGTAAACAGATTAGCAGAGCGAGGTATGTAGGCGGTGCTA 6709
DB 7007 CGCCACTGGCAGCAGCAGTGTGTAAACAGATTAGCAGAGCGAGGTATGTAGGCGGTGCTA 7066
QY 6710 CAGAGTTCTTGAGTGGTGGCTAACTACGCTCACTAGAGGACAGTATTTGCTATCT 6769
DB 7067 CAGAGTTCTTGAGTGGTGGCTAACTACGCTCACTAGAGGACAGTATTTGCTATCT 7126
QY 6770 GCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAGAGTTGGTAGTCTTGTATCCGGCAAC 6829

DB 7127 GCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGGTAGTCTTGTATCCGGCAAC 7186
QY 6830 APACCAACCGCTGTAGCGGTGTTTTTTTGTGTAAGCAGCAGAGATTACGCGCAGAAAA 6889
DB 7187 AAACCAACCGCTGTAGCGGTGTTTTTTTGTGTAAGCAGCAGAGATTACGCGCAGAAAA 7246
QY 6890 AAGGATCTCAAGAAAGATCCCTTTGATCTTTCTACCGGGTCTGACGCTCAGTGGAAACAA 6949
DB 7247 AAGGATCTCAAGAAAGATCCCTTTGATCTTTCTACCGGGTCTGACGCTCAGTGGAAACAA 7306
QY 6950 ACTCAGTTAAAGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTT 7009
DB 7307 ACTCAGTTAAAGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTT 7366
QY 7010 TAAATTAATAATGAAGTCTTAAATCAATCTAAAGTATATAGTAAAACTTGGTCTGACA 7069
DB 7367 TAAATTAATAATGAAGTCTTAAATCAATCTAAAGTATATAGTAAAACTTGGTCTGACA 7426
QY 7070 GTTACCAATGCTTAATCAGTAGGACCTATCTCAGCGATCTGTCTATTTGTTCTATCCA 7129
DB 7427 GTTACCAATGCTTAAATCAGTAGGACCTATCTCAGCGATCTGTCTATTTGTTCTATCCA 7486
QY 7130 TAGTTCCTGACTCCCGCTGCTAGATAACTACGATACGGAGGCTTTACCATCTGGCC 7189
DB 7487 TAGTTCCTGACTCCCGCTGCTAGATAACTACGATACGGAGGCTTTACCATCTGGCC 7546
QY 7190 CAGTCTGCAATGATACCGGAGACCCACGCTACCGCTCCAGATTTATCAGCAATAA 7249
DB 7547 CAGTCTGCAATGATACCGGAGACCCACGCTACCGCTCCAGATTTATCAGCAATAA 7606
QY 7250 ACCAGCCAGCGGAAGGGCGGAGAGTGTCTCGAACTTTATCCGCTCCATCC 7309
DB 7607 ACCAGCCAGCGGAAGGGCGGAGAGTGTCTCGAACTTTATCCGCTCCATCC 7666
QY 7310 AGTCTATTAAATGTTTCCCGGGAAGCTAGAGTAGTTTCGCGAGTTAATAGTTTCGCA 7369
DB 7667 AGTCTATTAAATGTTTCCCGGGAAGCTAGAGTAGTTTCGCGAGTTAATAGTTTCGCA 7726
QY 7370 AGTGTGTTGCCATTGCTGAGGAGTGTGTGTCAGCTCGCTCGTGTGTTGTTATGGCTTCA 7429
DB 7727 AGTGTGTTGCCATTGCTGAGGAGTGTGTGTCAGCTCGCTCGTGTGTTGTTATGGCTTCA 7786
QY 7430 TCAGCTCCCGTTCCTCCCAACGATCAAGCGGAGTTACATGATCCCGCATGTTGTGCAAAAAG 7489
DB 7787 TCAGCTCCCGTTCCTCCCAACGATCAAGCGGAGTTACATGATCCCGCATGTTGTGCAAAAAG 7846
QY 7490 CGGTTAGCTCTTTCGGTCTCCGATCGTTGTCAGAAAGTAAAGTTGGCCGAGTGTATCAC 7549
DB 7847 CGGTTAGCTCTTTCGGTCTCCGATCGTTGTCAGAAAGTAAAGTTGGCCGAGTGTATCAC 7906
QY 7550 TCATGTTATGGCAGCAGTGCATTAATCTCTTACTGTCTATGCCATCCGTAAGATGCTTT 7609
DB 7907 TCATGTTATGGCAGCAGTGCATTAATCTCTTACTGTCTATGCCATCCGTAAGATGCTTT 7966
QY 7610 CTGTGACTGGTGAAGTACTCAACCAAGTCAATCTGAGAAATAGTGTATGCGGCGACCGAGTT 7669
DB 7967 CTGTGACTGGTGAAGTACTCAACCAAGTCAATCTGAGAAATAGTGTATGCGGCGACCGAGTT 8026
QY 7670 GCTTTCGCGGCGTCAACCGGATAATACCGGCGCATAGCAGAACTTTAAAGTGC 7729
DB 8027 GCTTTCGCGGCGTCAACCGGATAATACCGGCGCATAGCAGAACTTTAAAGTGC 8086
QY 7730 TCATCAATGGAAGAGTCTTTCGGGCGGAAAACTCTCAAGGATCTTACCGCTGTTGAGAT 7789
DB 8087 TCATCAATGGAAGAGTCTTTCGGGCGGAAAACTCTCAAGGATCTTACCGCTGTTGAGAT 8146
QY 7790 CAGTTTCGATGTAACCACTCGTGACCCAACTGATCTTCAGCACTTTTATCTTACCA 7849
DB 8147 CAGTTTCGATGTAACCACTCGTGACCCAACTGATCTTCAGCACTTTTATCTTACCA 8206
QY 7850 GCGTTTTCGCTGAGCAAAAACAGGAAGTGCAGAAATGCGCAAAAAGGGAATAAGGCGCA 7909

Db	8207	GCCTTTCTGGGTGAGCAAAAACAGGAAGCAAATGCCGCAAAAAAAGGGAATAAGGGCGA	8266
QY	7910	CACGGAATGTTGAATACTCATCTCTTCTCTTTTTCAATATTATTGTAAGCATTATTATCAGG	7969
Db	8267	CACGGAATGTTGAATACTCATCTCTTCTCTTTTTCAATATTATTGTAAGCATTATTATCAGG	8326
QY	7970	GTTATTGTCTCATGACGGGATACATATTTGAATGTATTTAGAAAAATAACAATAAGGGG	8029
Db	8327	GTTATTGTCTCATGACGGGATACATATTTGAATGTATTTAGAAAAATAACAATAAGGGG	8386
QY	8030	TTCGGCGCACATTTCCCGGAAAAAGTGCCACTGACGCTCTAAGAACAACATTTATTATCATGA	8089
Db	8387	TTCGGCGCACATTTCCCGGAAAAAGTGCCACTGACGCTCTAAGAACAACATTTATTATCATGA	8446
QY	8090	CATTAACTTATAAATATGAGCGTATCACGAGGCCCCCTTCGTCTTCAAGAAATTAATTCTCA	8149
Db	8447	CATTAACTTATAAATATGAGCGTATCACGAGGCCCCCTTCGTCTTCAAGAAATTAATTCTCA	8506
QY	8150	TGTTGACAGCTTATCATCGATAAGCTGACTCATGTTGTTGTTGTAATGTAAGACGAGAT	8209
Db	8507	TGTTGACAGCTTATCATCGATAAGCTGACTCATGTTGTTGTTGTAATGTAAGACGAGAT	8566
QY	8210	CGGGAACACTGAAAAATAACAGTTATTATTTCG	8241
Db	8567	CGGGAACACTGAAAAATAACAGTTATTATTTCG	8598
RESULT 3	7925 bp DNA circular SYN 06-APR-2003		
AVI178045	Expression vector pPYPX88, complete sequence.		
LOCUS	AVI178045		
DEFINITION	Expression vector pPYPX88, complete sequence.		
ACCESSION	AVI178045		
VERSION	AVI178045.1 GI:29568868		
KEYWORDS	Expression vector pPYPX88		
SOURCE	Expression vector pPYPX88		
ORGANISM	artificial sequences; vectors.		
REFERENCE	1 (bases 1 to 7925)		
AUTHORS	Yao,Q., Peng,R. and Xiong,A.		
TITLE	Direct Submission		
JOURNAL	Submitted (12-NOV-2002) Shanghai Academy of Agricultural Sciences, Agro-Biotechnology Research Center, 2901 Beidi Rd, Shanghai 201106, China		
COMMENT	NCBI staff are still waiting for submitters to provide appropriate coding region information.		
FEATURES	Location/Qualifiers		
source	1..7925		
	/organism="Expression vector pPYPX88"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:218197"		
	/lab_host="Pichia pastoris"		
promoter	1..936		
	/note="modified modified AOX1 promoter"		
misc_feature	937..1212		
	/note="similar to modified alpha factor secretion signal"		
misc_feature	complement(1994..4527)		
	/note="similar to histidinol dehydrogenase"		
misc_feature	4773..5529		
	/note="3' AOX1 fragment from Pichia pastoris"		
rep_origin	complement(5937..6610)		
	/note="ColE1"		
CDS	complement(6755..7615)		
	/codon_start=1		
	/product="beta-lactamase"		
	/protein_id="BAO84025.1"		
	/db_xref="GI:29568863"		
	/translation="MSQHFRVALIPFFAARCLVPFAHPETLVKVKDAEDQLGARVGY VELNDISKILLESFRRPEPMNSTFKLLLGAVLIRVDAGCEOLGRIRHYSQNIDLYVE YSPYRKHTLDGMTRVELCSAAITMSDNTASNTALLLTITGGPKELTAFHNMGDHVRIRL DRWPELNEAIIPNDRDELTPPAAMATTKRLTLKLGLTSLARQQLIIDWMEADKKVGAVGPI LRSLAPGWFIADKSGAGERSGRIIALGDPKPSRIIVITVTYTTGSQTMDERNRQRTAF FIGASLIKHW"		

QY 1021 ACAACAGAGATGAACGCCACAAATCCGGCTGAAGCTGTATCGGTTACTCAGATTTA 1080
 DB 1012 ACTACCGAGGATGAACCTGCTCAAAATCCTGCTGAGGCTGTATCGGTTACTCAGCTG 1071
 QY 1081 GAAGGGATTTGCAATGTTGCTGTTTGGCAATTTTCCACAGACACAAATAACGGGTTATTG 1140
 DB 1072 GAGGTTGACTTCGAGCTGCTGTCTTGCCATTTCTTAATCCACCAACAGGTTTGTG 1131
 QY 1141 TTTTAAATACTACTATTATGCCAGCATTTCTGTTAAAGAAAGGGGTATCTCTCGAGAAA 1200
 DB 1132 -----GAGGAGCTGAAGCTGAAGCTGAACCTAAATTCATCAACACT 1173
 QY 1201 AGAGAGGCTGAAGCCAGGCCCGAGGCCAGGCCAGGCCAGGAAACATATGATCATG 1260
 DB 1174 ACTATCGCTTCTATCGCTGCTAAGGAGGAGGTTTCCCTCGAGAAAAGAGGCTGAG 1233
 QY 1261 GCCCGCCGGAGAGAGATAAAATTTGTGGCTTCCCGGTGTCAACCGCCAGCAGTGCACGAG 1320
 DB 1234 GCTTACGTAGAGTTGA-----1249
 QY 1321 AGAGTTGCTGTTTTTGATGACAGTGTCCGGGATTCCTGGTGTCTCCACCCCATGGCC 1380
 DB 1250 -----1249
 QY 1381 ATCGAGAACACTCAAGAGAGAAATGTCCCTTCTAATAGTGGCGTAGAATTCCTAGGG 1440
 DB 1250 -----GCTCGGG 1256
 QY 1441 CGCGCGGAATTAATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTGGGCACTTACG 1500
 DB 1257 CGCGCGGAATTAATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTGGGCACTTACG 1316
 QY 1501 AGAAGACCGGCTTCTAGATTTCTAATCAAGAGGATGTCAAGTGCCTATTTGCCCTGAG 1560
 DB 1317 AGAAGACCGGCTTCTAGATTTCTAATCAAGAGGATGTCAAGTGCCTATTTGCCCTGAG 1376
 QY 1561 ATGACGGCTTCATTTTGTATCTTTTATTTGTAACCTATATAGTATAGGATTTTATTT 1620
 DB 1377 ATGACGGCTTCATTTTGTATCTTTTATTTGTAACCTATATAGTATAGGATTTTATTT 1436
 QY 1621 GTCATTTGTTCTTCGTAGAGCTTCTCTGATCAGGCTATCTCGCAGCTGATGAA 1680
 DB 1437 GTCATTTGTTCTTCGTAGAGCTTCTCTGATCAGGCTATCTCGCAGCTGATGAA 1496
 QY 1681 TATCTTGTGTAGGGTTTGGAAATCATTCGAGTTGATGTTTCTTCTGTTATTTCCC 1740
 DB 1497 TATCTTGTGTAGGGTTTGGAAATCATTCGAGTTGATGTTTCTTCTGTTATTTCCC 1556
 QY 1741 ACTCCTCTTCAGAGTACAGAAATTAAGTGAAGTTCGTTTGTGCAAGCTTATCGATAA 1800
 DB 1557 ACTCCTCTTCAGAGTACAGAAATTAAGTGAAGTTCGTTTGTGCAAGCTTATCGATAA 1608
 QY 1801 GCTTTAATGCGGTAGTTTATCAGCTTAAATGCTTAAGCAGTCAAGCAGCAGCTGATGAA 1860
 DB 1609 GCTTTAATGCGGTAGTTTATCAGCTTAAATGCTTAAGCAGTCAAGCAGCAGCTGATGAA 1668
 QY 1861 ATCTAACAAATGCGCTCATCGTATCCTCGGCAACCTCAGCTGATGCTGATGAGTATAGG 1920
 DB 1669 ATCTAACAAATGCGCTCATCGTATCCTCGGCAACCTCAGCTGATGCTGATGAGTATAGG 1728
 QY 1921 CTTGGTTATGCGGTATCTGCGCGGCTTCTTGGGGATATCGTCCATTCGACAGCATGCG 1980
 DB 1729 CTTGGTTATGCGGTATCTGCGCGGCTTCTTGGGGATATCGTCCATTCGACAGCATGCG 1788
 QY 1981 CAGTCACTATGCGGTGCTAGCGTATATCGGTTGATGCAATTTCTATGCGCACCCGT 2040
 DB 1789 CAGTCACTATGCGGTGCTAGCGTATATCGGTTGATGCAATTTCTATGCGCACCCGT 1848
 QY 2041 TCTCGAGCACTGTCCGACCGCTTTGGCGCGCCGAGTCTGCTGCTTCTGCTACTTGG 2100
 DB 1849 TCTCGAGCACTGTCCGACCGCTTTGGCGCGCCGAGTCTGCTGCTTCTGCTACTTGG 1908
 QY 2101 AGCCACTATCGACTACGGCATCAATGGCGACACACACCCCTCTGTGTGATCTATCGAATCTA 2160

DB 1909 AGCCACTATCGACTACGGCATCATGCGACCAACCCGCTCTGTGATCTATCGAATCTTA 1968
 QY 2161 AATGTAAGTTAAATCTCTAAATAAATAAATTAAGTCCAGTTTCTCAATACGAACCTTAA 2220
 DB 1969 AATGTAAGTTAAATCTCTAAATAAATAAATTAAGTCCAGTTTCTCAATACGAACCTTAA 2028
 QY 2221 CAGCATTTGCGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAA 2280
 DB 2029 CAGCATTTGCGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAA 2088
 QY 2281 TATGTTTCAGTCCCTCAGGAGTTACGCTTGTGAAGTATGAACCTTCTGGAAGTTGCGAG 2340
 DB 2089 TATGTTTCAGTCCCTCAGGAGTTACGCTTGTGAAGTATGAACCTTCTGGAAGTTGCGAG 2148
 QY 2341 TGTAACTCCGCTGTATTGACGCGCATATCCGTTACGTTGGCAAGTGTGGTTGGTACCGG 2400
 DB 2149 TGTAACTCCGCTGTATTGACGCGCATATCCGTTACGTTGGCAAGTGTGGTTGGTACCGG 2208
 QY 2401 AGAGTAATCTCCACAACTCTCTGGAGAGTAGGCACCAACACAGATCCAGCGCTT 2460
 DB 2209 AGAGTAATCTCCACAACTCTCTGGAGAGTAGGCACCAACACAGATCCAGCGCTT 2268
 QY 2461 GTACTTTGATCAACATGAAGAAAGCAATCTCTGATTTGCAAGGATCAAGTCTTCAGGAGCGT 2520
 DB 2269 GTACTTTGATCAACATGAAGAAAGCAATCTCTGATTTGCAAGGATCAAGTCTTCAGGAGCGT 2328
 QY 2521 ACTGATGAGCAATTTCAAAGCCTCTCTGATGTTGCAACCGATAGGGTTGTAGAGTGTG 2580
 DB 2329 ACTGATGAGCAATTTCAAAGCCTCTCTGATGTTGCAACCGATAGGGTTGTAGAGTGTG 2388
 QY 2581 CAATACACTTGGCTGACAAATTTCAACCTTGGCAACTGCAAGCTTGGTTGTGAACAGCAT 2640
 DB 2389 CAATACACTTGGCTGACAAATTTCAACCTTGGCAACTGCAAGCTTGGTTGTGAACAGCAT 2448
 QY 2641 CTTCAATTTGCGAAGCTCTCTGCTCATATCGACAGCCAAAGAAATCACTCGGAAT 2700
 DB 2449 CTTCAATTTGCGAAGCTCTCTGCTCATATCGACAGCCAAAGAAATCACTCGGAAT 2508
 QY 2701 CAATACACTTGGCTGAGCAGAAAGCTCTGAGGCAACGAAATCTGATCAGCGTAT 2760
 DB 2509 CAATACACTTGGCTGAGCAGAAAGCTCTGAGGCAACGAAATCTGATCAGCGTAT 2567
 QY 2761 TATCAGCAATAAATTAGAACTTCAAGAGGCTTCAAGAGGCTTCAATCTACACAGGCTG 2820
 DB 2568 TATCAGCAATAAATTAGAACTTCAAGAGGCTTCAAGAGGCTTCAATCTACACAGGCTG 2627
 QY 2821 ATGTGCTATTTGAAACCAATCTTGGCAGCAGTAAAGAACTGTTTCTGGAACCAATA 2880
 DB 2628 ATGTGCTATTTGAAACCAATCTTGGCAGCAGTAAAGAACTGTTTCTGGAACCAATA 2687
 QY 2881 TTTTGTCAACTTAGGAAAGTCTTCTGCTTCCGTAAGCCATAGCAGCTACTGCTGGCGC 2940
 DB 2688 TTTTGTCAACTTAGGAAAGTCTTCTGCTTCCGTAAGCCATAGCAGCTACTGCTGGCGC 2747
 QY 2941 CTCTCTAGCAGCATACCTTAGCACCACCTTGTGGGCAACGATAGATGACTTCTGGGG 3000
 DB 2748 CTCTCTAGCAGCATACCTTAGCACCACCTTGTGGGCAACGATAGATGACTTCTGGGG 2807
 QY 3001 TRAGGGTACCATCTTCTTAGGTGAGATGCAAAAACAAATTTCTTGGCAACGAGCACTT 3060
 DB 2808 TRAGGGTACCATCTTCTTAGGTGAGATGCAAAAACAAATTTCTTGGCAACGAGCACTT 2867
 QY 3061 TGGCAGGAAACCCAGCATCAGGGAAGTGGAGGCAAGATTCGGTTCCACCGGAATAT 3120
 DB 2868 TGGCAGGAAACCCAGCATCAGGGAAGTGGAGGCAAGATTCGGTTCCACCGGAATAT 2927
 QY 3121 AGAGGCACTTTCTCAATAGTCTTGGCAACGAGAGCAGATACACCGGCAAGTCT 3180
 DB 2928 AGAGGCACTTTCTCAATAGTCTTGGCAACGAGAGCAGATACACCGGCAAGTCT 2987
 QY 3181 CAATCTGCAAGCTCTCGTTAGTGTAGCTTCAATGAAATTTCTGAGCTTATCTATAGAG 3240

Db 2988 CAACCTTGAACGCTCTCGCTTAGTGTAGCTTCAATGGAATTTCTCGAGCTTATCTATAGAGA 3047
 Qy 3241 GATCAATGGCTCTCTTAACGCTTATCTGGCAATTCGCAATGCTCTCTGGGAAAGGAGCTT 3300
 Db 3048 GATCAATGGCTCTCTTAACGCTTATCTGGCAATTCGCAATGCTCTCTGGGAAAGGAGCTT 3107
 Qy 3301 CTAACACAGGTGTCTTCAAGCGACTCCATCAAACTTGGCAGTTAGTTCTTAAAGGGCTT 3360
 Db 3108 CTAACACAGGTGTCTTCAAGCGACTCCATCAAACTTGGCAGTTAGTTCTTAAAGGGCTT 3167
 Qy 3361 TGTCAACATTTGACGACATTTGCGCAATTTGCTTGAATTAATTCATTAATCTGTTCCG 3420
 Db 3168 TGTCAACATTTGACGACATTTGCGCAATTTGCTTGAATTAATTCATTAATCTGTTCCG 3227
 Qy 3421 TTTTCTGTATAGGACGACGAGGCGATCTTCAATTTCTTGTGAGGAGGCGCTTAGAAACGT 3480
 Db 3228 TTTTCTGTATAGGACGACGAGGCGATCTTCAATTTCTTGTGAGGAGGCGCTTAGAAACGT 3287
 Qy 3481 CAATTTTGCACAAATTCATAGGACCTTCAGAGGACTTCTTTAGTTTGGATTTCTTCTT 3540
 Db 3288 CAATTTTGCACAAATTCATAGGACCTTCAGAGGAGCTTCTTTAGGTTTGGATTTCTTCTT 3347
 Qy 3541 TAGGTTGTCTTGGTGTATCTCGCTTGGCATCTCTCTTCTTCTAGTGACCTTTAGGG 3600
 Db 3348 TAGGTTGTCTTGGTGTATCTCGCTTGGCATCTCTCTTCTTCTAGTGACCTTTAGGG 3407
 Qy 3601 ACTTCATATCCAGGTTCTTCCAATGGTCCAAAGTCACACCGTACTTGGCACAICTAA 3660
 Db 3408 ACTTCATATCCAGGTTCTTCCAATGGTCCAAAGTCACACCGTACTTGGCACAICTAA 3467
 Qy 3661 CTAATGCAAAATTAATTAAGTCACACATTCGCCAGGCTATATCTTCTGCTGATTTAGCTT 3720
 Db 3468 CTAATGCAAAATTAATTAAGTCACACATTCGCCAGGCTATATCTTCTTCTGATTTAGCTT 3527
 Qy 3721 CTGCAAGTTCATCAGCTTCTCTCTTAAATTTAGGCTTC - AACAAAACTTCTGCTGCAAT 3779
 Db 3528 CTGCAAGTTCATCAGCTTCTCTCTTAAATTTAGGCTTC - AACAAAACTTCTGCTGCAAT 3587
 Qy 3780 AACGTTTGGTATAGAACCTTCGAGCATCTCTAGCATCCACACAGGTTGGCTTC 3839
 Db 3588 AACGTTTGGTATAGAACCTTCGAGCATCTCTAGCATCCACACAGGTT - GCTTCC 3646
 Qy 3840 ATGGCTCTAAGACCTTTGATTTGGCCAAACAGGAAGTGGCTTCCAAAGTGACAGAAACCA 3899
 Db 3647 ATGGCTCTAAGACCTTTGATTTGGCCAAACAGGAAGTGGCTTCCAAAGTGACAGAAACCA 3706
 Qy 3900 ACACCTGTTTGTTCACCAAAATTTCAAGCAGTCTCCATCAAACTCCAAATTCGATACCC 3959
 Db 3707 ACACCTGTTTGTTCACCAAAATTTCAAGCAGTCTCCATCAAACTCCAAATTCGATACCC 3766
 Qy 3960 AGCACTTTTGAAGTGTCTCCAGATGATACCTTTATACCAAAACCGTGACACAGAGAT 4019
 Db 3767 AGCACTTTTGAAGTGTCTCCAGATGATACCTTTATACCAAAACCGTGACACAGAGAT 3826
 Qy 4020 TGGTAGCTCCAGTTTGTCTTATAGCTTCGGAATFAGACTTTTGGAGGATACAC 4079
 Db 3827 TGGTAGCTCCAGTTTGTCTTATAGCTTCGGAATFAGACTTTTGGAGGATACAC 3886
 Qy 4080 AGGCCCAAGGATATTAAGAGGTCAGCCACCAAGAGTAGTAATAGACCATCGGGCGG 4139
 Db 3887 AGGCCCAAGGATATTAAGAGGTCAGCCACCAAGAGTAGTAATAGACCATCGGGCGG 3946
 Qy 4140 TCAGTAGTCAAGACGCGCAACAAATTTCACTGACAGGGAATTTTGAACATCTTCAGAA 4199
 Db 3947 TCAGTAGTCAAGACGCGCAACAAATTTCACTGACAGGGAATTTTGAACATCTTCAGAA 4006
 Qy 4200 AGTTGCTATTCAGTAGTCAATTCGCGAGCATCAATTAATGGGGATTTATACAGAGACACA 4259
 Db 4007 AGTTGCTATTCAGTAGTCAATTCGCGAGCATCAATTAATGGGGATTTATACAGAGACACA 4066
 Qy 4260 GTGGAAGTCACATCTACCAACTTTGCGGTCTCAGAAAAAGCATAAACAGATTCTACTACG 4319
 Db 4067 GTGGAAGTCACATCTACCAACTTTGCGGTCTCAGAAAAAGCATAAACAGATTCTACTACG 4126

Qy 4320 CCATTTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAGGCAAGGATCTACTAGCA 4379
 Db 4127 CCATTTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAGGCAAGGATCTACTAGCA 4186
 Qy 4380 TTAGCGGCAAGGATGCAACTTTTATCAACAGGCTCTATAGATAAACCTTAGCCCTTGGG 4439
 Db 4187 TTAGCGGCAAGGATGCAACTTTTATCAACAGGCTCTATAGATAAACCTTAGCCCTTGGG 4246
 Qy 4440 ATCATCTCTTTGGACAACTCTTTCTGCCAAATCTAGGTCCAAATCACTTCAATGATACCA 4499
 Db 4247 ATCATCTCTTTGGACAACTCTTTCTGCCAAATCTAGGTCCAAATCACTTCAATGATACCA 4306
 Qy 4500 TTATTTGACAACTTTGAGCAAGTTGTCGATCAGCTCTCAATTTGCTCTCTGTAACGGAT 4559
 Db 4307 TTATTTGACAACTTTGAGCAAGTTGTCGATCAGCTCTCAATTTGCTCTCTGTAACGGAT 4366
 Qy 4560 GACTCAACTTTGACAACTTTGAGCTCAGTTGAGTGAATGAGTGAATCTGATCAGGTTGTGC 4619
 Db 4367 GACTCAACTTTGACAACTTTGAGCTCAGTTGAGTGAATGAGTGAATCTGATCAGGTTGTGC 4426
 Qy 4620 AGCTGGTCAGCAAGCATAGGAAACACGCTTTTCTTACCAACTCAAGGAATTTATCAAC 4679
 Db 4427 AGCTGGTCAGCAAGCATAGGAAACACGCTTTTCTTACCAACTCAAGGAATTTATCAAC 4486
 Qy 4680 TCTGCAACACTTTGCGTATGCGAGTAGCAAGGAAATGTCTACTTTGAAGTCGGACAGTGA 4739
 Db 4487 TCTGCAACACTTTGCGTATGCGAGTAGCAAGGAAATGTCTACTTTGAAGTCGGACAGTGA 4546
 Qy 4740 GTGTAGTCTTGAGAAATTTCTGAAGCCGTTATTTTATTTATCATGAGTCAGTCATCAGGAG 4799
 Db 4547 GTGTAGTCTTGAGAAATTTCTGAAGCCGTTATTTTATTTATCATGAGTCAGTCATCAGGAG 4606
 Qy 4800 ATCTCTTACGCGGAGCGATGTCGCGACCTGCGAGGTCGGCATCACCGGCGGCGACAGGT 4859
 Db 4607 ATCTCTTACGCGGAGCGATGTCGCGACCTGCGAGGTCGGCATCACCGGCGGCGACAGGT 4639
 Qy 4860 GGGGTTGTCTGGCGCTATATCCGACATCACCGATGGGGAAGATCGGGCTCGGCACCTTC 4919
 Db 4640 GCTTGGGTGAACAGGCTTTTATTTCTGGCATCCACTAAATATAGCATGCCGC - 4689
 Qy 4920 GGGCTCATGAGCGCTTGTTCGGCGTGGTATGTTGGCAGCGCCCGTGGCGGGGAGCTG 4979
 Db 4690 - - - - - 4689
 Qy 4980 TTGGGCGGCATCTCCTTGTGATGCAACCATTTCTTGGCGGGGGTGTCTCAACGGGCTCAAC 5039
 Db 4690 - - - - - GGATTTCTTGGCGGGCGGTGCT - AACGGGCTCAAC 4724
 Qy 5040 CTACTACTGGGCTGTCTCTTAATCAGGAGTGCATTAAGGAGAGCGTCGAGTATCTATG 5099
 Db 4725 CTAC - GCGTGGCTGTCTCTTAATCAGGAGTGCATTAAGGAGAGCGTCGAGTATCTATG 4783
 Qy 5100 ATTTGGAAGTATCGGAATGCTGATACCCGATTTCTTCACTGTCTTGGAGTCTCTCTATCAGA 5159
 Db 4784 ATTTGGAAGTATCGGAATGCTGATACCCGATTTCTTCACTGTCTTGGAGTCTCTCTATCAGA 4843
 Qy 5160 TTATGCCCAACTTAAAGCAACCGGAGGAGGATTTTCATGTTGAATTTCTCTGACTTTTGG 5219
 Db 4844 TTATGCCCAACTTAAAGCAACCGGAGGAGGATTTTCATGTTGAATTTCTCTGACTTTTGG 4903
 Qy 5220 TCATCAGTAGTCTGGAATGCTGAGACTATCTCGGTATGACAGCAAGAAATGTCTCTTCTTG 5279
 Db 4904 TCATCAGTAGTCTGGAATGCTGAGACTATCTCGGTATGACAGCAAGAAATGTCTCTTCTTG 4963
 Qy 5280 GAGACAGTAAATGAAGTCCCAACCAATTAAGAAATCTTGTATATCAGGAACAAACTCTTCTTG 5339
 Db 4964 GAGACAGTAAATGAAGTCCCAACCAATTAAGAAATCTTGTATATCAGGAACAAACTCTTCTTG 5023
 Qy 5340 TTTTCGAACTTTTTCGGTGCCTTGAACATTAATTAAGTGTAGTGGATATGTCTGGGTAGGAAT 5399
 Db 5024 TTTTCGAACTTTTTCGGTGCCTTGAACATTAATTAAGTGTAGTGGATATGTCTGGGTAGGAAT 5083

5400 GGAGCGGCAAAATGCTTACCTTCTGGACCTTCAAGAGGTATGTAGGGTTTGTAGATACCTG 5459 QY
5084 GGAGCGGCAAAATGCTTACCTTCTGGACCTTCAAGAGGTATGTAGGGTTTGTAGATACCTG 5143 Db
5460 ATGCCAACTTCAGTGAACAAGTGTCTATTTGGTTCAAAACCAATCCGAAATCCAGAGAAATC 5519 QY
5144 ATGCCAACTTCAGTGAACAAGTGTCTATTTGGTTCAAAACCAATCCGAAATCCAGAGAAATC 5203 Db
5520 AAAGTTGTGTCTACTATTTGATCCAAAGCCAGTCCGGTCTTGAAATCTGACAAATAGTGTGC 5579 QY
5204 AAAGTTGTGTCTACTATTTGATCCAAAGCCAGTCCGGTCTTGAAATCTGACAAATAGTGTGC 5263 Db
5580 TCGTGTGTGTAGGTCATCTTTGTATGAATAAATCTAGTCTTTGTGATCTAAATATCTTTGAC 5639 QY
5264 TCGTGTGTGTAGGTCATCTTTGTATGAATAAATCTAGTCTTTGTGATCTAAATATCTTTGAC 5323 Db
5640 GAGCAAGCGGATAAATATACCAAACTCTTTTAAACCTTTTAAACCTTTAAAGGCAAGTAT 5699 QY
5324 GAGCAAGCGGATAAATATACCAAACTCTTTTAAACCTTTTAAACCTTTAAAGGCAAGTAT 5383 Db
5700 GTCTGCCGTGATTTAAACCCCAAAATCAGCTCTGACTCTGATCTCATCACTTGAGGGGCA 5759 QY
5384 GTCTGCCGTGATTTAAACCCCAAAATCAGCTCTGACTCTGATCTCATCACTTGAGGGGCA 5443 Db
5760 CTATCTCTTTTGTAGAGAAATTTGGGAGATCGGATATCGAGAAAAAGGTACGCTGATTTT 5819 QY
5444 CTATCTCTTTTGTAGAGAAATTTGGGAGATCGGATATCGAGAAAAAGGTACGCTGATTTT 5503 Db
5820 AAAAGTGAATTTATCTCAAGATCTCTGCCCTCGCGGCTTTCGGTGTATGACGGTCAAAACC 5879 QY
5504 AAAAGTGAATTTATCTCAAGATCTCTGCCCTCGCGGCTTTCGGTGTATGACGGTCAAAACC 5563 Db
5880 TCTGACACATGACGCTCCCGAGACGGTCAAGCTTTGCTGTAAAGCGGATCCCGGAGCA 5939 QY
5564 TCTGACACATGACGCTCCCGAGACGGTCAAGCTTTGCTGTAAAGCGGATCCCGGAGCA 5623 Db
5940 GACAAAGCCGTCAGGCGCGTCAGCGGTGTGGCGGGTGTGGGCGCGAGCCATGACCC 5999 QY
5624 GACAAAGCCGTCAGGCGCGTCAGCGGTGTGGCGGGTGTGGGCGCGAGCCATGACCC 5683 Db
6000 AGTCAGCTGAGTACGAGCGGTATCTGCTTAACTATCGGCATCAGAGCAGATTGT 6059 QY
5684 AGTCAGCTGAGTACGAGCGGTATCTGCTTAACTATCGGCATCAGAGCAGATTGT 5743 Db
6060 ACTGAGAGTGACCAATATGCGGTGTGAATACCGCACAGATCGGTAAAGGAAATACCG 6119 QY
5744 ACTGAGAGTGACCAATATGCGGTGTGAATACCGCACAGATCGGTAAAGGAAATACCG 5803 Db
6120 CATCAGGCGCTTCCGCTTCTCCTCTCACTGACTCGCTCGCTCGGTTCGGCTGCG 6179 QY
5804 CATCAGGCGCTTCCGCTTCTCCTCTCACTGACTCGCTCGCTCGGTTCGGCTGCG 5863 Db
6180 GCGAGCGGTATCAGTCACTCAAGCGGTATATACGGTTATCCAGAAATCAGGGGATAA 6239 QY
5864 GCGAGCGGTATCAGTCACTCAAGCGGTATATACGGTTATCCAGAAATCAGGGGATAA 5923 Db
6240 CGCAGGAAAGAAATGTAGCAAAAGGCGAGCAAAAGGCGAGAACCGTAAAGGCGCG 6299 QY
5924 CGCAGGAAAGAAATGTAGCAAAAGGCGAGCAAAAGGCGAGAACCGTAAAGGCGCG 5983 Db
6300 GTTGTGCGGTTTTTCATAGCTCCGCCCTCGAGCATCAAAAATTCGACGCTC 6359 QY
5984 GTTGTGCGGTTTTTCATAGCTCCGCCCTCGAGCATCAAAAATTCGACGCTC 6043 Db
6360 AAGTCAGAGTGGCGAAACCCGACAGGACTATAAGATACAGGCGTTTCCCGCTGGAAG 6419 QY
6044 AAGTCAGAGTGGCGAAACCCGACAGGACTATAAGATACAGGCGTTTCCCGCTGGAAG 6103 Db
6420 CTCCTCTGTGCGCTCTCTCTGTTCCGACCCCTGCGGTACCGATACCTGTCGCGCTTCT 6479 QY
6104 CTCCTCTGTGCGCTCTCTCTGTTCCGACCCCTGCGGTACCGATACCTGTCGCGCTTCT 6163 Db
6480 CCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCAAGCTGTAGGTATCTCAGTTCCGTTGTA 6539 QY

6164 CCCTTCGGGAAGCGTGGCGCTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCCGTTGTA 6223 Db
6540 GGTGTGTGCTCCAAAGCTGGGCTGTGTGCAAGAAACCCCGCTTTCAGCCGAGCCGCTGGC 6599 QY
6224 GGTGTGTGCTCCAAAGCTGGGCTGTGTGCAAGAAACCCCGCTTTCAGCCGAGCCGCTGGC 6283 Db
6600 CTTATCCGTAATACTATGCTTTCAGTCCAAACCCGCTAAGACACGACTTATCCGCACTGCG 6659 QY
6284 CTTATCCGTAATACTATGCTTTCAGTCCAAACCCGCTAAGACACGACTTATCCGCACTGCG 6343 Db
6660 AGCAGCACTGTAATACTGATTCAGAGAGGAGTATGTAGGCGGTGTACAGAGTTCTT 6719 QY
6344 AGCAGCACTGTAATACTGATTCAGAGAGGAGTATGTAGGCGGTGTACAGAGTTCTT 6403 Db
6720 GAAGTGTGGCTTAACCTACGCTACACTAGAAGGACAGTATTTGCTATCTCGCTCTGCT 6779 QY
6404 GAAGTGTGGCTTAACCTACGCTACACTAGAAGGACAGTATTTGCTATCTCGCTCTGCT 6463 Db
6780 GAAGCAGTATACCTTCGGAAGAGAGTGTGATCCGGCAAAACCAACCCG 6839 QY
6464 GAAGCAGTATACCTTCGGAAGAGAGTGTGATCCGGCAAAACCAACCCG 6523 Db
6840 TGGTAGCGGTGTGTTTTGTTGCAAGCAGAGATACCGCAGAAAAAGGATCTCA 6899 QY
6524 TGGTAGCGGTGTGTTTTGTTGCAAGCAGAGATACCGCAGAAAAAGGATCTCA 6583 Db
6900 AGAAGATCTTTGATCTTTTCTACGGGTCTGACGCTCAAGTGAAGCAAACTCAGTTA 6959 QY
6584 AGAAGATCTTTGATCTTTTCTACGGGTCTGACGCTCAAGTGAAGCAAACTCAGTTA 6643 Db
6960 AGGATTTTGGTCATGAGATTTCAAAAAGGATCTTCACTAGATCTTTTAAATTTAA 7019 QY
6644 AGGATTTTGGTCATGAGATTTCAAAAAGGATCTTCACTAGATCTTTTAAATTTAA 6703 Db
7020 ATCAAGTTTTTAAATCAATCTAAAGTATATATAGTAAACTTGTCTGACAGTTACCAATG 7079 QY
6704 ATGAGTTTTTAAATCAATCTAAAGTATATATAGTAAACTTGTCTGACAGTTACCAATG 6763 Db
7080 CTTAATCAGTAGGCACTATCTCAGCGATCTGCTATTTTGGTTTCAATCAGTTGCTG 7139 QY
6764 CTTAATCAGTAGGCACTATCTCAGCGATCTGCTATTTTGGTTTCAATCAGTTGCTG 6823 Db
7140 ACTCCCGCTGTAGTAACTACGATACGGAGGCTTACCAATCTGGGCCCAAGTCTGC 7199 QY
6824 ACTCCCGCTGTAGTAACTACGATACGGAGGCTTACCAATCTGGGCCCAAGTCTGC 6883 Db
7200 AATGATACCGGAGACCCAGCTCACCGCTCCAGATTTATCAGCAATAAACCCAGCCAGC 7259 QY
6884 AATGATACCGGAGACCCAGCTCACCGCTCCAGATTTATCAGCAATAAACCCAGCCAGC 6943 Db
7260 CGGAAGGCGCGAGCGAGAGTGTCTGCACTTTATCGGCTCCATCCAGTCTATTAA 7319 QY
6944 CGGAAGGCGCGAGCGAGAGTGTCTGCACTTTATCGGCTCCATCCAGTCTATTAA 7003 Db
7320 TTGTTGCGGGAAGCTAGATAGTTTCCGCAAGTTAATAGTTTGGCAACCGTTGTC 7379 QY
7004 TTGTTGCGGGAAGCTAGATAGTTTCCGCAAGTTAATAGTTTGGCAACCGTTGTC 7063 Db
7380 CATTTGTCAGGCACTCGTGTGTCAAGCTGTGTTGGTATGGCTTCATTCAGTCCGG 7439 QY
7064 CATTTGTCAGGCACTCGTGTGTCAAGCTGTGTTGGTATGGCTTCATTCAGTCCGG 7123 Db
7440 TTCCCAACGATCAAGCGAGTTACATGATCCCATGTTGTGCAAAAAAGCGTTAGCTC 7499 QY
7124 TTCCCAACGATCAAGCGAGTTACATGATCCCATGTTGTGCAAAAAAGCGTTAGCTC 7183 Db
7500 CTTCCGCTCTCCGATCGTTCTCAGAAAGTAAAGTTGGCGCAGTGTATCACTCATGTTAT 7559 QY
7184 CTTCCGCTCTCCGATCGTTCTCAGAAAGTAAAGTTGGCGCAGTGTATCACTCATGTTAT 7243 Db
7560 GCGAGCACTGCAATAATTTCTTACTGTCAATCGCATCCGTAAGATGCTTTCTGTGACTGG 7619 QY

Db 7244 GCGACGACGTCATATCTCTTCTCTGTCATGCCATCGTAAGATGCTTTTCTGTGACTGG 7303
Qy 7620 TGAGTACTCAACCAAGTCATCTCTGAGAAATAGTGTATGCGGCGACCGAGTGTCTCTTGGCCC 7679
Db 7304 TGAGTACTCAACCAAGTCATCTCTGAGAAATAGTGTATGCGGCGACCGAGTGTCTCTTGGCCC 7363
Qy 7680 GCGGTCACACGCGATATACCGCGCCACATAGCAGAACTTTAAAGTGCTCATCTATGG 7739
Db 7364 GCGGTCACACGCGATATACCGCGCCACATAGCAGAACTTTAAAGTGCTCATCTATGG 7423
Qy 7740 AAAAGCTTCTTGGGGGGAAGAACTCTCAAGGATCTTACCGCTGTGTGAGATCCAGTTCGAT 7799
Db 7424 AAAAGCTTCTTGGGGGGAAGAACTCTCAAGGATCTTACCGCTGTGTGAGATCCAGTTCGAT 7483
Qy 7800 GTAAACCACTCGTGACCCCACTGATCTTTCAGCATCTTTTACTTTTACCCAGCGTTCCTGG 7859
Db 7484 GTAAACCACTCGTGACCCCACTGATCTTTCAGCATCTTTTACTTTTACCCAGCGTTCCTGG 7543
Qy 7860 GTGACCAAAAACAGGAGCGCAAAATGCCGCAAAAAGGGAATAGGGCGCACACGGAATG 7919
Db 7544 GTGACCAAAAACAGGAGCGCAAAATGCCGCAAAAAGGGAATAGGGCGCACACGGAATG 7603
Qy 7920 TTGAATCTCATCTCTTCTTTCATATATTGAAGCAATTTATCAGGGTTATTTGCT 7979
Db 7604 TTGAATCTCATCTCTTCTTTCATATATTGAAGCAATTTATCAGGGTTATTTGCT 7663
Qy 7980 CATGAGCGGATACATATTGTAATGTTTGAATAATAGAAAAATAAACAATAGGGGTTCCGCGCAC 8039
Db 7664 CATGAGCGGATACATATTGTAATGTTTGAATAATAGAAAAATAAACAATAGGGGTTCCGCGCAC 7723
Qy 8040 ATTTCCCGGAAAGTGCCACCTGACGCTTAAGAAACCATTTATATCATGACATTAACCTA 8099
Db 7724 ATTTCCCGGAAAGTGCCACCTGACGCTTAAGAAACCATTTATATCATGACATTAACCTA 7783
Qy 8100 TAAAAATAGGCGTATCAGAGCGCCTTTCGCTTCAAGAAATATTTCTCATGTTTGACAG 8159
Db 7784 TAAAAATAGGCGTATCAGAGCGCCTTTCGCTTCAAGAAATATTTCTCATGTTTGACAG 7843
Qy 8160 CTTATCATCGATAAGCTGACTCATGTTGGTATTTGAAATAGACGCGAGATCGGGAACACT 8219
Db 7844 CTTATCATCGATAAGCTGACTCATGTTGGTATTTGAAATAGACGCGAGATCGGGAACACT 7903
Qy 8220 GAAAAATACAGTTATTATTCG 8241
Db 7904 GAAAAATACAGTTATTATTCG 7925

RESULT 4

LOCUS PPPIC9 8020 bp DNA linear PLN 22-OCT-1994
DEFINITION P.pastoris DNA for pPIC9 expression vector.
ACCESSION Z46233
VERSION Z46233.1 GI:559514
KEYWORDS beta-lactamase.
SOURCE Pichia pastoris
ORGANISM Pichia pastoris
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Pichia.
REFERENCE 1 (bases 1 to 7957)
AUTHORS Sreerishna,K.
Scorer,C.A., Clare,J.J., McCombie,W.R., Romanos,M.A. and
Rapid selection using G418 of high copy number transformants of
Pichia pastoris for high-level foreign gene expression
Biotechnology (N.Y.) 12 (2), 181-184 (1994)
TITLE
JOURNAL 94153511
MEDLINE 7764433
PUBMED
REFERENCE 2 (bases 1 to 8020)
AUTHORS Romanos,M.A.
TITLE Direct Submission
JOURNAL Submitted (18-OCT-1994) Romanos M. A., Wellcome Research
Laboratories, Biotechnology Division, Langley Court, Beckenham Kent
TN4

source
1..8020
/organism="Pichia pastoris"
/mol_type="genomic DNA"
/db_xref="taxon:4922"
/clone="pPIC3"
1..8020
/standard_name="pPIC9"
/function="Pichia alpha-factor secretion vector"
/label=pPIC9
1..354
/standard_name="AOX1"
complement (1250..3280)
/standard_name="HIS4"
/codon_start=1
/protein_id="CAA86304.1"
/db_xref="GI:559515"
/db_xref="GOA:Q12201"
/db_xref="SPTREMBL:Q12201"
/translation="MTFFLPAYASVAEFDNLSLVGKAVFPYAADQLHNLKFTQST
ELQNVQVESSTVDQFELIDNLLKLYNGINEVILDLDLAERVQVQMIPGARVIYR
TLIDKVASLPANASTADVPSPPLGDLKSFNGSGRTVYAFSTAKLVDTVSTVASGII
PLIDAROLTTVEVLSADVKKPVPSEILLASLTDRPDGLFTTLVADSNVSLGVSS
KKSPEARTGTGYQSRHGLWYKGTSGTOKLGLDLDGDCGLKVVVEONGVGF
CHLERTSCFGSKGLRANAEFCGIVRAMLKVLIPNGVLTTFCLNKKIREADELAE
AKSEKPEVPSSEGRILECKIDVSKASSEIEDALRRPIQTEQIMELVKPIVDNVQ
NGDKALLBELTAFDFGVALKTVLEAPFPELMQLPDNVKRAIDLSIDNVKHFHAQLT
ETQVETCPGVVCSRFAPIEKVGLYIPGGTALPSTSLMLGVPAKVAGCKEIVFASP
PKQDGLTPEVIYVAHKVAKCIVLAGGAQAAVAMAYGTETVPKCDKI FPGNGQVTA
AKMVQNTSALCSIDMPAGPSEVIVIAADKVDPDFVASDLLLLKLVNVLIPR"
3589..4036
/standard_name="3' AOX1"
complement (5606..6466)
/product="beta-lactamase"
/standard_name="AmpR"
5766..7703
/standard_name="AOX1"
7727..7979
/standard_name="alpha factor leader"

ORIGIN

Query Match 81.6%; Score 6726.6; DB 8; Length 8020;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 6784; Conservative 0; Mismatches 4; Indels 15; Gaps 4;
Qy 1440 GCGGCGCGAATTAATTCGCTTACAGATGACTGTTCCTCAGTTCAGTTCGGGACCTTAC 1499
Db 1 GCGGCGCGAATTAATTCGCTTACAGATGACTGTTCCTCAGTTCAGTTCGGGACCTTAC 60
Qy 1500 GAGAGACCGGCTTGTAGATTCTTAATCAAGAGATGTCAGATGCCATTGCTGTAGA 1559
Db 61 GAGAGACCGGCTTGTAGATTCTTAATCAAGAGATGTCAGATGCCATTGCTGTAGA 120
Qy 1560 GATCAGGCTTCATTTTGTATCTTTTATTTGTAACTATATAGTATAGATTTTTC 1619
Db 121 GATCAGGCTTCATTTTGTATCTTTTATTTGTAACTATATAGTATAGATTTTTC 180
Qy 1620 TGTCAATTTGTTCTTCTCGTACGAGTTCCTCGATCAGCCTATCTCGCAGCTGTAGA 1679
Db 181 TGTCAATTTGTTCTTCTCGTACGAGTTCCTCGATCAGCCTATCTCGCAGCTGTAGA 240
Qy 1680 ATATCTTGTGTAGGGGTTTGGGAAATCATTCAGATTGTATGTTTCTTGTATTTC 1739
Db 241 ATATCTTGTGTAGGGGTTTGGGAAATCATTCAGATTGTATGTTTCTTGTATTTC 300
Qy 1740 GACTCTCTTTCAGATACAGAAATTAAGTACAGAGTTCGTTTGTGCAAGCTTATCGATA 1799
Db 301 GACTCTCTTTCAGATACAGAAATTAAGTACAGAGTTCGTTTGTGCAAGCTTATCGATA 360
Qy 1800 AGCTTTAATCGGTAGTTTATCACAGTTTAAATTCGTAACGAGTCAGCACCCTGTATGA 1859

1860 AATCTAAATGCGCTCATCGTCTCCCGCAACCGTCAACCTCGATGCTGTAGGCATAG 1919
421 AATCTAAATGCGCTCATCGTCTCCCGCAACCGTCAACCTCGATGCTGTAGGCATAG 480
1920 GCTTGGTATGCGCGTATGCGCGGCTCTTGGCGGATATCGTCAATTCGCAAGATCG 1979
481 GCTTGGTATGCGCGTATGCGCGGCTCTTGGCGGATATCGTCAATTCGCAAGATCG 540
1980 CCAGTCATATGCGGCTGTCTAGCGCTATATGCGTTCGATCAATTTCTATGCGCACCG 2039
541 CCAGTCATATGCGGCTGTCTAGCGCTATATGCGTTCGATCAATTTCTATGCGCACCG 600
2040 TTCTCGGAGCACTGTCGACCGCTTTGGCCGCGCCAGTCTCTGCTCGCTTCGCTACTTG 2099
601 TTCTCGGAGCACTGTCGACCGCTTTGGCCGCGCCAGTCTCTGCTCGCTTCGCTACTTG 660
2100 GAGCCACTATCGACTACCGGATCATGCGGACCAACCGCTCTGCGATCTATCGAATCT 2159
661 GAGCCACTATCGACTACCGGATCATGCGGACCAACCGCTCTGCGATCTATCGAATCT 720
2160 AAATGTAAGTTAAATCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2219
721 AAATGTAAGTTAAATCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 780
2220 ACAGCAATTCGCGTGAAGCATATAGACCTTCAACGAGCGGACGATCGATCACTGCTTGGCA 2279
781 ACAGCAATTCGCGTGAAGCATATAGACCTTCAACGAGCGGACGATCGATCACTGCTTGGCA 840
2280 ATATGTTTCAGTCCCTCAGGAGTTACGCTTGTGAAGTGAATCACTTCTGGAAGTTGCA 2339
841 ATATGTTTCAGTCCCTCAGGAGTTACGCTTGTGAAGTGAATCACTTCTGGAAGTTGCA 900
2340 GTGTTAACTCCGCTATATGAGCGGATATCGTACGTTGCAAGTGTGTTGCTACCG 2399
901 GTGTTAACTCCGCTATATGAGCGGATATCGTACGTTGCAAGTGTGTTGCTACCG 960
2400 GAGGAGTAATTCACCACTCTCTGAGGATAGGACCAACCAACAGATCCAGCGGT 2459
961 GAGGAGTAATTCACCACTCTCTGAGGATAGGACCAACCAACAGATCCAGCGGT 1020
2460 TGTAATTCATCAACATAAGAGAAGCAATTCGATTTGCAAGATCAAGTGTTCAGGAGCG 2519
1021 TGTAATTCATCAACATAAGAGAAGCAATTCGATTTGCAAGATCAAGTGTTCAGGAGCG 1080
2520 TACTGATGAGCATTTCCAAAGCTCTGCTAGGTTGCAACCGATAGGTTGTAGAGTGT 2579
1081 TACTGATGAGCATTTCCAAAGCTCTGCTAGGTTGCAACCGATAGGTTGTAGAGTGT 1140
2580 GCAATACACTTCGCTACCAATTTCAACCTTGGCAACTGCAAGCTTGGTGTGAACAGCA 2639
1141 GCAATACACTTCGCTACCAATTTCAACCTTGGCAACTGCAAGCTTGGTGTGAACAGCA 1200
2640 TCTTCAATTCGCGAAGCTCTTGTCTGTCATATCGACGCAACAGAAATCACCTGGGAA 2699
1201 TCTTCAATTCGCGAAGCTCTTGTCTGTCATATCGACGCAACAGAAATCACCTGGGAA 1260
2700 TCAATACCATGTTCAAGTTGACAGAGAAGTCTGAGGCAACGAAATCTGGATCAGCGTAT 2759
1261 TCAATACCATGTTCAAGTTGAGCTTGAAGCTCTGAGGCAACGAAATCTGGATCAGCGTAT 1319
2760 TTATCAGCAATAACTAGAACTTCAGAGGCCCGCAGGCGATGTCAATACTACAGAGGCT 2819
1320 TTATCAGCAATAACTAGAACTTCAGAGGCCCGCAGGCGATGTCAATACTACAGAGGCT 1379
2820 GATGTGCTATTTGAACCATCATCTTGGCAGCAGTAAAGAACTGTTTCTTGGACCAAT 2879
1380 GATGTGCTATTTGAACCATCATCTTGGCAGCAGTAAAGAACTGTTTCTTGGACCAAT 1439
2880 ATTTTGTACACTTAGGAACAGTTTCTGTTCCGTAAGCATAGCAGTACTGCTTGGGCG 2939
1440 ATTTTGTACACTTAGGAACAGTTTCTGTTCCGTAAGCATAGCAGTACTGCTTGGGCG 1499
2940 CCTCTGCTAGCACGATACATTAACCAACCTTGTGGGCAACGATGATGATCTCTGGG 2999

1500 CCTCTGCTAGCACGATACACCTAGCACCAACCTTGTGGGCAACGATGATGACTTCTGGG 1559
3000 GTAGGGTACCATCTCTTCTAGGTGGAGATGAAACAAATTTCTTTGCAACGACAACT 3059
1560 GTAGGGTACCATCTCTTCTAGGTGGAGATGAAACAAATTTCTTTGCAACGACAACT 1619
3060 TTGCGCAGGAACACCCAGCATCAGGGAAGTGGGAAGGAGAAATTCGCGTCCACGAGAAAT 3119
1620 TTGCGCAGGAACACCCAGCATCAGGGAAGTGGGAAGGAGAAATTCGCGTCCACGAGAAAT 1679
3120 TAGAGGCCAATCTTCTCAATAGGCTTTGCAAAACGAGAGACACTACACAGGGCAAGTC 3179
1680 TAGAGGCCAATCTTCTCAATAGGCTTTGCAAAACGAGAGACACTACACAGGGCAAGTC 1739
3180 TCAACTTGAACGCTCTCGGTAGTCTCATGGAATTTCTGACGTTATCTATAGAG 3239
1740 TCAACTTGAACGCTCTCGGTAGTCTCATGGAATTTCTGACGTTATCTATAGAG 1799
3240 AGATCAATAGGCTCTTAAAGCTTATCTGGAATTCGCAATTAAGTTCCTCTGGGAAAGAGCT 3299
1800 AGATCAATAGGCTCTTAAAGCTTATCTGGAATTCGCAATTAAGTTCCTCTGGGAAAGAGCT 1859
3300 TCTAAACAGAGGTCTTCAAGCGACTCCATCAAACTTCGCGAGTTAGTTCTTAAAGGGCT 3359
1860 TCTAAACAGAGGTCTTCAAGCGACTCCATCAAACTTCGCGAGTTAGTTCTTAAAGGGCT 1919
3360 TTGTCACCAATTTGACGAACATTTGTCGACAAATTTGTTGACAAATTCATTAATCTCTTC 3419
1920 TTGTCACCAATTTGACGAACATTTGTCGACAAATTTGTTGACAAATTCATTAATCTCTTC 1979
3420 GTTTTCTGATAGGACGACGAGGCGATCTTCAATTTCTTGTGAGGAGGCTTTAGAAAG 3479
1980 GTTTTCTGATAGGACGACGAGGCGATCTTCAATTTCTTGTGAGGAGGCTTTAGAAAG 2039
3480 TCAATTTTGACAAATTTCAATACGACCTTCAGAGGAGCTTCTTTAGGTTTGAATCTTCT 3539
2040 TCAATTTTGACAAATTTCAATACGACCTTCAGAGGAGCTTCTTTAGGTTTGAATCTTCT 2099
3540 TTAGGTTTCTTGTGTTATCTGCGTTGGCATCTCTTCTCTCTAGTACCTTTAGG 3599
2100 TTAGGTTTCTTGTGTTATCTGCGTTGGCATCTCTTCTCTCTAGTACCTTTAGG 2159
3600 GACTTCATATCCAGGTTCTCTCCACTCGTCCAAAGTCAACCGTACTTGGCAGACATCTA 3659
2160 GATTCATATCCAGGTTCTCTCCACTCGTCCAAAGTCAACCGTACTTGGCAGACATCTA 2219
3660 ACTAATGCAAAATATAAATAAGTCAGCACATTTCCAGGCTATATCTTCTTGGATTTAGCT 3719
2220 ACTAATGCAAAATATAAATAAGTCAGCACATTTCCAGGCTATATCTTCTTGGATTTAGCT 2279
3720 TCTGCAAGTTTCAGGTTCTCTCTAAATTTTAGCGTTC-AACAAATCTCGTCTGCA 3778
2280 TCTGCAAGTTTCAGGTTCTCTCTAAATTTTAGCGTTC-AACAAATCTCGTCTGCA 2339
3779 TAAACCGTTTGGTATAAGAACCTTCTGAGCATTTGCTTACGATCCCAAGGTTGGCTTC 3838
2340 TAAACCGTTTGGTATAAGAACCTTCTGAGCATTTGCTTACGATCCCAAGGTTGGCTTC 2398
3839 CATGGCTCTAAGACCTTTGATTTGGCCAAACAGGAAGTGGCTTCAAGTACAGAAACC 3898
2399 CATGGCTCTAAGACCTTTGATTTGGCCAAACAGGAAGTGGCTTCAAGTACAGAAACC 2458
3899 AACACCTGTTTGTTCACCAACAAATTTCAAGAGTCTCCATCAATTCGAATTCGATACC 3958
2459 AACACCTGTTTGTTCACCAACAAATTTCAAGAGTCTCCATCAATTCGAATTCGATACC 2518
3959 CAGCAACTTTTGTAGTTGCTCCAGATGATGACCTTTTATACCAAAACCGTACAGAGAG 4018
2519 CAGCAACTTTTGTAGTTGCTCCAGATGATGACCTTTTATACCAAAACCGTACAGAGAG 2578
4019 TTGGTAGACTCCAGTTTGTGCTTATAGCTCCGGAATAGACTTTTGGACGAGTACAC 4078

2579 TTGTTAGACTCCAGTTTGTGTCTTATAGGCTCCGGAATAGACTTTTGTGACGAGTACAC 2638
Db
4079 CAGGCCCAACGAGTAATTAGAAGAGTCAAGCCCAAAAGTAGTAGTAATAGACCAATCGGGGCG 4138
Qy
2639 CAGGCCCAACGAGTAATTAGAAGAGTCAAGCCCAAAAGTAGTAGTAATAGACCAATCGGGGCG 2698
Db
4139 GTCACTAGTCAAGAGCCAAACAAATTTCTGACAGGGAACTTTTGTGACATCTTCTCAGA 4198
Qy
2699 GTCACTAGTCAAGAGCCAAACAAATTTCTGACAGGGAACTTTTGTGACATCTTCTCAGA 2758
Db
4199 AAGTTCTGATTCAGTATGCTCAATTCGCGAGCATCAATATGGGATATATACCAAGAACAC 4258
Qy
2759 AAGTTCTGATTCAGTATGCTCAATTCGCGAGCATCAATATGGGATATATACCAAGAACAC 2818
Db
4259 AGTGAAGTCACTCTACCAACTTTTGGGCTCTCAGAAAAGCAATPAAACAGTTCTACTACC 4318
Qy
2819 AGTGAAGTCACTCTACCAACTTTTGGGCTCTCAGAAAAGCAATPAAACAGTTCTACTACC 2878
Db
4319 GCCATTAGTGAACCTTTTCAATCGCCAGTGGAGAGAAAGGCAACGGATACCTAGC 4378
Qy
2879 GCCATTAGTGAACCTTTTCAATCGCCAGTGGAGAGAAAGGCAACGGATACCTAGC 2938
Db
4379 ATTAGCGGCAAGGATGCAACTTTTATCAACAGGCTCTCTATAGATAACCTTAGCGCTGG 4438
Qy
2939 ATTAGCGGCAAGGATGCAACTTTTATCAACAGGCTCTCTATAGATAACCTTAGCGCTGG 2998
Db
4439 GATCATCTTTGGCAACTCTTTCTGCGCAATCTAGTCCAAATCACTTCACTATGATACC 4498
Qy
2999 GATCATCTTTGGCAACTCTTTCTGCGCAATCTAGTCCAAATCACTTCACTATGATACC 3058
Db
4499 ATTATTGACAACTTGACAGAGTTGTGATCAGCTCTCTCAAAATGGTCTCTGTAAAGGA 4558
Qy
3059 ATTATTGACAACTTGACAGAGTTGTGATCAGCTCTCTCAAAATGGTCTCTGTAAAGGA 3118
Db
4559 TGACTCAACTTGCAATPAACTTGAAGCTCAGTCAATGAGTGAACCTGATCAGTTGTG 4618
Qy
3119 TGACTCAACTTGCAATPAACTTGAAGCTCAGTCAATGAGTGAACCTGATCAGTTGTG 3178
Db
4619 CAGCTGGTCAAGAGTGAAGAAACAGCGTTTCTTACCAAACTCAAGGAAATTTATCAAA 4678
Qy
3179 CAGCTGGTCAAGAGTGAAGAAACAGCGTTTCTTACCAAACTCAAGGAAATTTATCAAA 3238
Db
4679 CTCTGCAACACTTGCGTATGAGTGAAGGAAATGTCTATCTTGAAGTTCGACAGTG 4738
Qy
3239 CTCTGCAACACTTGCGTATGAGTGAAGGAAATGTCTATCTTGAAGTTCGACAGTG 3298
Db
4739 AGTGTAGTCTTGAGAAATCTGAGCGGTATTTTATATCAGTCACTCAGTCACTCAGGA 4798
Qy
3299 AGTGTAGTCTTGAGAAATCTGAGCGGTATTTTATATCAGTCACTCAGTCACTCAGGA 3358
Db
4799 GATCTCTACGCGGAGCATCTGTGGCGAAGTGCAGGTCGCGCATCACCGCGCCACAGG 4858
Qy
3359 GATCTCTACGCGGAGCATCTGTGGCGAAGTGCAGGTCGCGCATCACCGCGCCACAGG 3406
Db
4859 TGCGGTTGTGGCGCTATATCGCGGAGCATCAACGATGGGGAAGATCGGCTCGCCACTT 4918
Qy
3407 TGCGGTTGTGGCGCTATATCGCGGAGCATCACCGATGGGGAAGATCGGCTCGCCACTT 3466
Db
4919 CGGGCTCATGAGCGCTTGTCTTGGCGTGGTATGTTGGCGAGCCCGCTGGCGGGGACT 4978
Qy
3467 CGGGCTCATGAGCGCTTGTCTTGGCGTGGTATGTTGGCGAGCCCGCTGGCGGGGACT 3526
Db
4979 GTTGGCGCCATCTCTTTCATGCAACCAATCTCTTGGCGCGCGGTGTCTCAACGGCTCAA 5038
Qy
3527 GTTGGCGCCATCTCTTTCATGCAACCAATCTCTTGGCGCGCGGTGTCTCAACGGCTCAA 3586
Db
5039 CTTACTCTGGGCTGCTTCTTAAATGAGGAGTTCGATGAGGAGGAGGCTCGAGTATCTAT 5098
Qy
3587 CTTACTCTGGGCTGCTTCTTAAATGAGGAGTTCGATGAGGAGGAGGCTCGAGTATCTAT 3646
Db
5099 GATTGGAAGTATGGAATGGTATACCCGCAATCTTCTAGTCTCTTGGAGTCTCTATACAG 5158
Qy
3647 GATTGGAAGTATGGAATGGTATACCCGCAATCTTCTAGTCTCTTGGAGTCTCTATACAG 3706
Db

5159 ATTATGCCCAACTAAAGCAACCGGAGGAGAGATTTCTATGTAATTTCTCTGACTTTTG 5218
Qy
3707 ATTATGCCCAACTAAAGCAACCGGAGGAGAGATTTCTATGTAATTTCTCTGACTTTTG 3766
Db
5219 GTTCATCAGTACTCGAACTGTGAGACTATCTCGGTTATGACAGAGAAATGTCTCTCTT 5278
Qy
3767 GTTCATCAGTACTCGAACTGTGAGACTATCTCGGTTATGACAGAGAAATGTCTCTCTT 3826
Db
5279 GGAGACAGTAATGAAGTCCCAACCAATAAAGAAATCTTGTATTAGGAAACAACTTCTT 5338
Qy
3827 GGAGACAGTAATGAAGTCCCAACCAATAAAGAAATCTTGTATTAGGAAACAACTTCTT 3886
Db
5339 GTTTCGAACCTTTTGGTGGCTTGAACCTATAAATGATAGTGGATATGTCGGGTAGGAA 5398
Qy
3887 GTTTCGAACCTTTTGGTGGCTTGAACCTATAAATGATAGTGGATATGTCGGGTAGGAA 3946
Db
5399 TGGAGCGGCGCAAAATGCTTACCTTCTGGAACCTTCAAGAGGTATGTAGGTTTGTAGACT 5458
Qy
3947 TGGAGCGGCGCAAAATGCTTACCTTCTGGAACCTTCAAGAGGTATGTAGGTTTGTAGACT 4006
Db
5459 GATGCCAACTTCAGTGACAAAGTTGCTATTTTCTGATCCAAACCATTCGGAATCCAGAGAAAT 5518
Qy
4007 GATGCCAACTTCAGTGACAAAGTTGCTATTTTCTGATCCAAACCATTCGGAATCCAGAGAAAT 4066
Db
5519 CAAAGTTGTTGCTTACTATTGATCCAAAGCAGTGGCTTCTTGAACCTGACAAATAGTGTG 5578
Qy
4067 CAAAGTTGTTGCTTACTATTGATCCAAAGCAGTGGCTTCTTGAACCTGACAAATAGTGTG 4126
Db
5579 CTCGTGTTTGTAGTCACTCTTGTATGAATAAATCTAGTCTTGTATCTAAATATCTTGA 5638
Qy
4127 CTCGTGTTTGTAGTCACTCTTGTATGAATAAATCTAGTCTTGTATCTAAATATCTTGA 4186
Db
5639 CGAGCAAGCGGATATAATACCAAACTTAAAACTCTTTTAAACCTTAAAGGCAAGTA 5698
Qy
4187 CGAGCAAGCGGATATAATACCAAACTTAAACCTTAAACCTTAAAGGCAAGTA 4246
Db
5699 TGTCTGCTGTATTAACCCCAATCAGTCTGATCTGATCTCATCACTTGAAGGCGG 5758
Qy
4247 TGTCTGCTGTATTAACCCCAATCAGTCTGATCTGATCTCATCACTTGAAGGCGG 4306
Db
5759 ACTATCTGTTTGTAGAGAAATTTGCGAGATGCGATATCGAGAAAGAGTACGCTGATT 5818
Qy
4307 ACTATCTGTTTGTAGAGAAATTTGCGAGATGCGATATCGAGAAAGAGTACGCTGATT 4366
Db
5819 TAAAGCTGAAATTTTATCTCAAGATCTCTGCTCGCGCTTTCGCTGATGACGCTGAAAC 5878
Qy
4367 TAAAGCTGAAATTTTATCTCAAGATCTCTGCTCGCGCTTTCGCTGATGACGCTGAAAC 4426
Db
5879 CTCTGACATCAGCTCTCGGAGAGCTCAGCTCTCTGTAAGCGGATGCGGAGC 5938
Qy
4427 CTCTGACATCAGCTCTCGGAGAGCTCAGCTCTCTGTAAGCGGATGCGGAGC 4486
Db
5939 AGACAAGCCCGTCAGGCGGCTGTCAGCGGTGTGCGGGGTGTGCGGGCGCAGCATGACC 5998
Qy
4487 AGACAAGCCCGTCAGGCGGCTGTCAGCGGTGTGCGGGGTGTGCGGGCGCAGCATGACC 4546
Db
5999 CAGTCACTGAGTGAAGTGAATGCTGCTTAACTATGCGGCATCAGAGCAGATTG 6058
Qy
4547 CAGTCACTGAGTGAAGTGAATGCTGCTTAACTATGCGGCATCAGAGCAGATTG 4606
Db
6059 TACTGAGTGCACCATATGCGGTGTAATAATACCCACAGATGCGTAAGGAGAAATACC 6118
Qy
4607 TACTGAGTGCACCATATGCGGTGTAATAATACCCACAGATGCGTAAGGAGAAATACC 4666
Db
6119 GCATCAGGCGCTCTTCCGCTTCCCTCGCTCACTGATCTGCTGCTGCTGCTGCTGCTGCTG 6178
Qy
4667 GCATCAGGCGCTCTTCCGCTTCCCTCGCTCACTGATCTGCTGCTGCTGCTGCTGCTGCTG 4726
Db
6179 GCGGAGCGGTATCAGTCACTCAAAAGCGGTAAATACGGTTATCCACAGATCAGGAGATA 6238
Qy
4727 GCGGAGCGGTATCAGTCACTCAAAAGCGGTAAATACGGTTATCCACAGATCAGGAGATA 4786
Db

QY 6239 ACGCAGGAAGAACATGTGAGCAAAAGCCAGCAAAAGCCAGGAACCGTAAAGAGCCG 6298
DB 4787 ACGCAGGAAGAACATGTGAGCAAAAGCCAGCAAAAGCCAGGAACCGTAAAGAGCCG 4846
QY 6299 CGTTGTCGCGGTTTTTCCATAGGCTCCGCCGCCCTCGAGAGCATCACAAAATCGAGCCT 6358
DB 4847 CGTTGTCGCGGTTTTTCCATAGGCTCCGCCGCCCTCGAGAGCATCACAAAATCGAGCCT 4906
QY 6359 CAAATCAGAGGTGGGAAACCCGACAGAGCATTAAGATACAGAGGTTTCCCTCGGAA 6418
DB 4907 CAAATCAGAGGTGGGAAACCCGACAGAGCATTAAGATACAGAGGTTTCCCTCGGAA 4966
QY 6419 GCTCCCTCGTGGCTCTCTGTTCCGACCTCGCTCCGCTTACCGGATACCTGTCGCGCTTTC 6478
DB 4967 GCTCCCTCGTGGCTCTCTGTTCCGACCTCGCTTACCGGATACCTGTCGCGCTTTC 5026
QY 6479 TCCCTTCGGAAGAGGTGGGCTTTTCTCATAGTCAAGCTGAGGTATCTCATGTTGCGTGT 6538
DB 5027 TCCCTTCGGAAGAGGTGGGCTTTTCTCATAGTCAAGCTGAGGTATCTCATGTTGCGTGT 5086
QY 6539 AGGTGCTTCGCTCCAAAGCTGGGCTGTGACGACGACCCCGTTACGAGCCGACGCTGCG 6598
DB 5087 AGGTGCTTCGCTCCAAAGCTGGGCTGTGACGACGACCCCGTTACGAGCCGACGCTGCG 5146
QY 6599 CTTTATCCGGTAACTATCGTTGAGTCCAAACCCGGTAAAGACAGACTTATCGCCACTGG 6658
DB 5147 CTTTATCCGGTAACTATCGTTGAGTCCAAACCCGGTAAAGACAGACTTATCGCCACTGG 5206
QY 6659 CAGCAGCCACTGGTAAACAGGATACAGAGCAGGATATGAGCGGCTGCTACAGATTCT 6718
DB 5207 CAGCAGCCACTGGTAAACAGGATACAGAGCAGGATATGAGCGGCTGCTACAGATTCT 5266
QY 6719 TGAAGTGGTGGCTAACTACGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGC 6778
DB 5267 TGAAGTGGTGGCTAACTACGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGC 5326
QY 6779 TGAAGCAGATTACCTTCGGAAGAGGTTGAGTCTTGATTCGGGCAAAACAAACCCAGC 6838
DB 5327 TGAAGCAGATTACCTTCGGAAGAGGTTGAGTCTTGATTCGGGCAAAACAAACCCAGC 5386
QY 6839 CTGCTAGCGGTGGTTTTTTGTTTCAAGCAGCAGATACCGGCAAGAAAAGAGATCTC 6898
DB 5387 CTGCTAGCGGTGGTTTTTTGTTTCAAGCAGCAGATACCGGCAAGAAAAGAGATCTC 5446
QY 6899 AAGAAGATCTTTGATCTTTCTACGGGTCTGACGCTCAGTGGAAAGAACTCAGCTT 6958
DB 5447 AAGAAGATCTTTGATCTTTCTACGGGTCTGACGCTCAGTGGAAAGAACTCAGCTT 5506
QY 6959 AAGGATTTTGTCTAGGATTAACAAAAGATCTTCACTAGATCTTTTAAATTAA 7018
DB 5507 AAGGATTTTGTCTAGGATTAACAAAAGATCTTCACTAGATCTTTTAAATTAA 5566
QY 7019 AATGAAGTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAAT 7078
DB 5567 AATGAAGTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAAT 5626
QY 7079 GCTTAAATCAGTGGGACCTATCTCAGGATCTGCTATTTGTTCTATCCATAGTTGCTT 7138
DB 5627 GCTTAAATCAGTGGGACCTATCTCAGGATCTGCTATTTGTTCTATCCATAGTTGCTT 5686
QY 7139 GACTCCCGTCTGTAGATAACTACGATACGGAGGGCTTACCATCTGCGCCCGAGTGTG 7198
DB 5687 GACTCCCGTCTGTAGATAACTACGATACGGAGGGCTTACCATCTGCGCCCGAGTGTG 5746
QY 7199 CAATGATACCGGACCCACGCTCACCGGCTCCAGATTTATCAAGCAATAAACACCGCAG 7258
DB 5747 CAATGATACCGGACCCACGCTCACCGGCTCCAGATTTATCAAGCAATAAACACCGCAG 5806
QY 7259 CCGGAGGGCCGAGCGAGAGTGTCTGCACTTTATCCGCTCCCTCAGTCTATTA 7318
DB 5807 CCGGAGGGCCGAGCGAGAGTGTCTGCACTTTATCCGCTCCCTCAGTCTATTA 5866
QY 7319 ATTGTTGCGGGAAGCTAGAGTAAAGTGTGCGCAAGTTAATAGTTTGGCAACGTTGTTG 7378

DB 5867 ATTGTTGCGGGAAGCTAGAGTAAAGTTCGCGAGTTAATAGTTTGCAGCAAGTTGTTG 5926
QY 7379 CCATTGCTGAGGAGCATCGTGTCTCAGCTCGTCTGTTGTTGTTGTTGTTGTTGTTGTTG 7438
DB 5927 CCATTGCTGAGGAGCATCGTGTCTCAGCTCGTCTGTTGTTGTTGTTGTTGTTGTTGTTG 5986
QY 7439 GTTCCCAACGATCAAGCGAGATTACATGATCCCCCAATGTTGTTGTTGTTGTTGTTGTTG 7498
DB 5987 GTTCCCAACGATCAAGCGAGTTACATGATCCCCCAATGTTGTTGTTGTTGTTGTTGTTG 6046
QY 7499 CTTTCGCTCTCCGATCGTTGTCAGAGTAAGTTGCGCGCAGTGTGTTATCACTCATGTTA 7558
DB 6047 CTTTCGCTCTCCGATCGTTGTCAGAGTAAGTTGCGCGCAGTGTGTTATCACTCATGTTA 6106
QY 7559 TGGCAGCAGCTGCATAAATCTCTTACTGTCATGTCATCCGTAAGATGCTTTTCTGTTGACTG 7618
DB 6107 TGGCAGCAGCTGCATAAATCTCTTACTGTCATGTCATCCGTAAGATGCTTTTCTGTTGACTG 6166
QY 7619 GTGAGTACTCAACCAAGTCAATCTGAGAAATAGTGTATGCGGCGACCGAGTGTCTTGGC 7678
DB 6167 GTGAGTACTCAACCAAGTCAATCTGAGAAATAGTGTATGCGGCGACCGAGTGTCTTGGC 6226
QY 7679 CGGCGTCAACCAAGTCAATCTGAGAAATAGTGTATGCGGCGACCGAGTGTCTTGGC 7738
DB 6227 CGGCGTCAACCAAGTCAATCTGAGAAATAGTGTATGCGGCGACCGAGTGTCTTGGC 6286
QY 7739 GAAACGTTCTTTCGGGCGGAAACTCTCAGAGTCTTACCGCTGTTGAGATCCAGTTCGA 7798
DB 6287 GAAACGTTCTTTCGGGCGGAAACTCTCAGAGTCTTACCGCTGTTGAGATCCAGTTCGA 6346
QY 7799 TGTAAACCCACTCTGTCAGCCCAACTGATCTTTCAGCATCTTTTACTTTTACCAGCGTTTCTG 7858
DB 6347 TGTAAACCCACTCTGTCAGCCCAACTGATCTTTCAGCATCTTTTACTTTTACCAGCGTTTCTG 6406
QY 7859 GGTGAGCAAAAACAGGAGCAAAATGCGGCAAAAGGGAATAAGGCGGACACGGAAT 7918
DB 6407 GGTGAGCAAAAACAGGAGCAAAATGCGGCAAAAGGGAATAAGGCGGACACGGAAT 6466
QY 7919 GTTGAATACCTACTACTCTTCTCTTCTTCAATATTATTGAAGCATTTATCAGGGTTATTGTC 7978
DB 6467 GTTGAATACCTACTACTCTTCTCTTCTTCAATATTATTGAAGCATTTATCAGGGTTATTGTC 6526
QY 7979 TCATAGCGGATACATATTGTAATGTTATTAGAAAAATAACAAATAGGGGTTCCGCGCA 8038
DB 6527 TCATAGCGGATACATATTGTAATGTTATTAGAAAAATAACAAATAGGGGTTCCGCGCA 6586
QY 8039 CATTTCCCGAAAGTGCCACCTGAGCTCTAGAAACCATTATTATCATGACATTAACT 8098
DB 6587 CATTTCCCGAAAGTGCCACCTGAGCTCTAGAAACCATTATTATCATGACATTAACT 6646
QY 8099 ATAAAAATAGCGGTATCAGAGGCCCTTTTCGCTCTTCAAGAAATTAATTCATGTTGACA 8158
DB 6647 ATAAAAATAGCGGTATCAGAGGCCCTTTTCGCTCTTCAAGAAATTAATTCATGTTGACA 6706
QY 8159 GCTTATCATGATAGCTGACTCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 8218
DB 6707 GCTTATCATGATAGCTGACTCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 6766
QY 8219 TGAATAAATACAGTTATTATTCG 8241
DB 6767 TGAATAAATACAGTTATTATTCG 6789

RESULT 5
PPPIC3
LOCUS
DEFINITION P.pastoris DNA for pPIC3 expression vector.
ACCESSION Z46231
VERSION Z46231.1
KEYWORDS beta-lactamase.
SOURCE Pichia pastoris
ORGANISM Pichia pastoris

PPPIC3 7757 bp DNA linear
PLN 22-OCT-1994

Mon Sep 27 14:45:38 2004

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Pichia.
 1 (bases 1 to 7757)
 Scorer, C.A., Clare, J.J., McCombie, W.R., Romanos, M.A. and
 Sreekrishna, K.
 Rapid selection using G418 of high copy number transformants of
 Pichia pastoris for high-level foreign gene expression
 Biotechnology (N.Y.) 12 (2), 181-184 (1994)
 94153511
 7764433
 2 (bases 1 to 7757)
 Romanos, M.A.
 Direct Submission
 Submitted (18-OCT-1994) Romanos M. A., Wellcome Research
 Laboratories, Biology Division, Langley Court, Beckenham, Kent, UK,
 BR3 3BS
 Location/Qualifiers
 1. 7757
 /organism="Pichia pastoris"
 /mol_type="genomic DNA"
 /db_xref="taxon:4922"
 /clone="pPIC3"
 1. 7757
 /standard_name="pPIC3"
 /function="Pichia intracellular expression vector"
 1. 354
 /standard_name="AOX1"
 complement(1237..3267)
 /standard_name="HIS4"
 3589..4036
 /standard_name="3' AOX1"
 complement(5606..6466)
 /product="beta-lactamase"
 /standard_name="AmpR"
 6766..7703
 /standard_name="AOX1"
 ORIGIN
 Query Match 81.5%; Score 6713.6; DB 8; Length 7757;
 Best Local Similarity 99.7%; Pred No. 0;
 Matches 6771; Conservative 0; Mismatches 4; Indels 15; Gaps 4;
 1453 AATTCGCCTTAGACATGACTGTCTCTCAGTTCAAGTTGGGCACTTACGAGAAGACCGGTC 1512
 1 AATTCGCCTTAGACATGACTGTCTCTCAGTTCAAGTTGGGCACTTACGAGAAGACCGGTC 60
 1513 TTGCTAGATCTTAATCAAGAGGATGTCAGATGCCATTTGCTGAGAGATCAGGCTTCA 1572
 61 TTGCTAGATCTTAATCAAGAGGATGTCAGATGCCATTTGCTGAGAGATCAGGCTTCA 120
 1573 TTTTGTGATCTTTTATTTGTAACCTATATAGTATAGGATTTTGTGCAATTTGTT 1632
 121 TTTTGTGATCTTTTATTTGTAACCTATATAGTATAGGATTTTGTGCAATTTGTT 180
 1633 CTTCTGATGAGCTTGTCTGATCAGCCTATCTCGAGCTGATGAATATCTTGTGTA 1692
 181 CTTCTGATGAGCTTGTCTGATCAGCCTATCTCGAGCTGATGAATATCTTGTGTA 240
 1693 GGGGTTTGGGAAATCATTCGAGTTTGTATGTTTCTTGGTATTTCCACCTCTTTCAG 1752
 241 GGGGTTTGGGAAATCATTCGAGTTTGTATGTTTCTTGGTATTTCCACCTCTTTCAG 300
 1753 AGTACAGAGATTAAGTGAAGATTCGTTTGTGAGCTTATCGATAGCTTTAATGCGG 1812
 301 AGTACAGAGATTAAGTGAAGATTCGTTTGTGAGCTTATCGATAGCTTTAATGCGG 360
 1813 TAGTTTATCACAGTTAAATTTGTAACGCACTGAGCAGCAGCTGATGAAATCTTAAATGC 1872
 361 TAGTTTATCACAGTTAAATTTGTAACGCACTGAGCAGCAGCTGATGAAATCTTAAATGC 420
 1873 GGTATGTCATCTCGGCAACCGTCAACCTGATGCTGAGCATAGCTTGGTTGGTATGCC 1932
 421 GGTATGTCATCTCGGCAACCGTCAACCTGATGCTGAGCATAGCTTGGTTGGTATGCC 480

1933 GGTACTGCGGGGCTCTTTGGGGATATCGTCCATTCGACAGCATCGCCAGTCACTATGG 1992
 481 GGTACTGCGGGGCTCTTTGGGGATATCGTCCATTCGACAGCATCGCCAGTCACTATGG 540
 1993 COTGCTGTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCGCTTCTCGAGCACT 2052
 541 COTGCTGTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCGCTTCTCGAGCACT 600
 2053 GTCCGACCGCTTTGCGCGCCGCCAGTCTCTCGCTTCGCTTCTGCTTCTGAGGCCATATCGA 2112
 601 GTCCGACCGCTTTGCGCGCCGCCAGTCTCTCGCTTCGCTTCTGCTTCTGAGGCCATATCGA 660
 2113 CTACCGCATGATGGCGACCAACACCGCTCTCTGAGATCTATCGAATCTAAATGTAAGTTAA 2172
 661 CTACCGCATGATGGCGACCAACACCGCTCTCTGAGATCTATCGAATCTAAATGTAAGTTAA 720
 2173 AATCTCTAAATTAATTAATTAAGTCCAGTTCTCCATACGAACCTTAACAGCATTCGGGT 2232
 721 AATCTCTAAATTAATTAATTAAGTCCAGTTCTCCATACGAACCTTAACAGCATTCGGGT 780
 2233 GAGCATCTAGACCTTCAACAGCAGCAGATCCATCCTGCTTGGCCAAATATGTTTCAGTC 2292
 781 GAGCATCTAGACCTTCAACAGCAGCAGATCCATCCTGCTTGGCCAAATATGTTTCAGTC 840
 2293 CTTCAAGGATTTACGTTCTTGTGAAGTGAATGAATCTTGAAGGTTGCAAGTGTAACTCCGC 2352
 841 CTTCAAGGATTTACGTTCTTGTGAAGTGAATGAATCTTGAAGGTTGCAAGTGTAACTCCGC 900
 2353 TGTATTGACGGGCATATCCGTACGTTGGCAAGTGTGTTGTTGTAACGGAGGTAATCTC 2412
 901 TGTATTGACGGGCATATCCGTACGTTGGCAAGTGTGTTGTTGTAACGGAGGTAATCTC 960
 2413 CACAACTCTCTGAGAGTAGGACCAACAAACACAGATCCAGCGTGTGTACTTGATCAA 2472
 961 CACAACTCTCTGAGAGTAGGACCAACAAACACAGATCCAGCGTGTGTACTTGATCAA 1020
 2473 CATAGAGAAGCATTTCTGATTTCCAGATCAAGTGTTCAGGAGCTACTGATTGGACA 2532
 1021 CATAGAGAAGCATTTCTGATTTCCAGATCAAGTGTTCAGGAGCTACTGATTGGACA 1080
 2533 TTTTCCAAAGCCCTGCTGAGTTGCAACCGATAGGTTGCAACCGATAGGTTGTAGAGTGTGCAATACATTGC 2592
 1081 TTTTCCAAAGCCCTGCTGAGTTGCAACCGATAGGTTGTAGAGTGTGCAATACATTGC 1140
 2593 GTACAAATTTCAACCCCTTGGCACTGACAGCTTGTGTTGTAACAGCATCTTCAATTTCTGG 2652
 1141 GTACAAATTTCAACCCCTTGGCACTGACAGCTTGTGTTGTAACAGCATCTTCAATTTCTGG 1200
 2653 CAAGCTCTTGTCTGTCATATCGACAGCCAAACAGATCACTTGGGAATCAATPACCATGTT 2712
 1201 CAAGCTCTTGTCTGTCATATCGACAGCCAAACAGATCACTTGGGAATCAATPACCATGTT 1260
 2713 CAGCTTTGAGACAGAGGTCTGAGGCAACGAAATCTGGATCAGCGTATTTATCAGCAATAA 2772
 1261 CAGCTTTGAG - CAGAGGCTCTGAGGCAACGAAATCTGGATCAGCGTATTTATCAGCAATAA 1319
 2773 CTAGAACTTTCAGAAAGCCCGACGAGCATGCTCAATACTACACAGGGCTGATGTTGTCATTTT 2832
 1320 CTAGAACTTTCAGAAAGCCCGACGAGCATGCTCAATACTACACAGGGCTGATGTTGTCATTTT 1379
 2833 GAACCATCATCTTGGCAGCAGTAAAGAACTGGTTTCTCGAACAAATATTTTGTACACT 2892
 1380 GAACCATCATCTTGGCAGCAGTAAAGAACTGGTTTCTCGAACAAATATTTTGTACACT 1439
 2893 TAGGAACAGTTTCTGTTCCGTAGCCATAGCAGTACTCTGCGGCGCTCTCTGCTAGCA 2952
 1440 TAGGAACAGTTTCTGTTCCGTAGCCATAGCAGTACTCTGCGGCGCTCTCTGCTAGCA 1499
 2953 CGATACACTTAGCAACCAACTTGTGGGCAACGATAGTATGATGATGATGATGATGATGATGAT 3012
 1500 CGATACACTTAGCAACCAACTTGTGGGCAACGATAGTATGATGATGATGATGATGATGATGAT 1559

QY 3013 CCTTCTTAGTGGATGCAAAAAAATTTCTTTGCAACCGAGCAACTTTGGCAGAACAC 3072
DB 1560 CTTTCTTAGTGGATGCAAAAAAATTTCTTTGCAACCGAGCAACTTTGGCAGAACAC 1619
QY 3073 CCAGCATCAGGGAAGTGGAGGAGCAAAATTTGGGTTCCACAGGAGATATAGAGGCCAACTT 3132
DB 1620 CCAGCATCAGGGAAGTGGAGGAGCAAAATTTGGGTTCCACAGGAGATATAGAGGCCAACTT 1679
QY 3133 TCTCAATAGGTCTTGCAAAAAAAGAGAGCAGATACACAGGGAAGTCTCAACTTGCACAG 3192
DB 1680 TCTCAATAGGTCTTGCAAAAAAAGAGAGCAGATACACAGGGAAGTCTCAACTTGCACAG 1739
QY 3193 TCTCCGTAGTGGATGCAAGTATTTCTTGAGGATTTCTTGAGGATTTCTATAGAGATCAATGGCTC 3252
DB 1740 TCTCCGTAGTGGATGCAAGTATTTCTTGAGGATTTCTTGAGGATTTCTATAGAGATCAATGGCTC 1799
QY 3253 TCTTAAAGCTTATCTGCAATTTGATAGTTCCTCTGGGAAAGGAGCTTCTAACACAGGTG 3312
DB 1800 TCTTAAAGCTTATCTGCAATTTGATAGTTCCTCTGGGAAAGGAGCTTCTAACACAGGTG 1859
QY 3313 TCTTCAAGAGCTTCAATCAACTTTGGGAGTATTTCTTAAAGGGCTTTGTACACATTTT 3372
DB 1860 TCTTCAAGAGCTTCAATCAACTTTGGGAGTATTTCTTAAAGGGCTTTGTACACATTTT 1919
QY 3373 GACGAACATTTGCGAATTTGGTCTTGAATTTGACTAATTTCCATTAATCTGTCTCCGTTTCTGGATAG 3432
DB 1920 GACGAACATTTGCGAATTTGGTCTTGAATTTGACTAATTTCCATTAATCTGTCTCCGTTTCTGGATAG 1979
QY 3433 GACGAGAGGCTTCTCAATTTCTTGAGGAGGCTTGAAGAAOGFCAATTTTGACACA 3492
DB 1980 GACGAGAGGCTTCTCAATTTCTTGAGGAGGCTTGAAGAAOGFCAATTTTGACACA 2039
QY 3493 ATTCAATACGACTTTCAGAGGAGCTTCTTGAGGTTTGGATTTCTTTTGGTGTGTTCCCT 3552
DB 2040 ATTCAATACGACTTTCAGAGGAGCTTCTTGAGGTTTGGATTTCTTTTGGTGTGTTCCCT 2099
QY 3553 TGGGTATCTCGGCTTGGCATCTCTCTTCTTCTAGTGACCTTTAGGAGCTTCTATATCCA 3612
DB 2100 TGGGTATCTCGGCTTGGCATCTCTCTTCTTCTAGTGACCTTTAGGAGCTTCTATATCCA 2159
QY 3613 GGTCTTCTCCACCTCGTCCAGCTCAGACCGTACTTGGGACATCTAATCAATGCAAAAT 3672
DB 2160 GGTCTTCTCCACCTCGTCCAGCTCAGACCGTACTTGGGACATCTAATCAATGCAAAAT 2219
QY 3673 AAAATAAGTCAGCATAATTTCCAGGCTATATCTTCTTGGATTTAGCTTCTGCAAGTTTCAT 3732
DB 2220 AAAATAAGTCAGCATAATTTCCAGGCTATATCTTCTTGGATTTAGCTTCTGCAAGTTTCAT 2279
QY 3733 CAGCTTCTCCCTAATTTTAGGTTTC-AACAAAATTTCTGCTGCAAAATAACCGTTTGGTA 3791
DB 2280 CAGCTTCTCCCTAATTTTAGGTTTCACAAAATTTCTGCTGCAAAATAACCGTTTGGTA 2339
QY 3792 TAAGAACCTTCTTGAGGATTTGCTCTTACGATCCACAGGTTGCTTCCATGGCTCTAAGA 3851
DB 2340 TAAGAACCTTCTTGAGGATTTGCTCTTACGATCCACAGGTTGCTTCCATGGCTCTAAGA 2398
QY 3852 CCTTTGATTTGGCCTTAAACAGGAAGTGGTTTCCAAAGTGACAGAAACCAACCTGTTTGT 3911
DB 2399 CCTTTGATTTGGCCTTAAACAGGAAGTGGTTTCCAAAGTGACAGAAACCAACCTGTTTGT 2458
QY 3912 TCAACCAAAATTTCAAGAGCTTCCATCAATCCCAATTTGATACCCAGCAACTTTTGA 3971
DB 2459 TCAACCAAAATTTCAAGAGCTTCCATCAATCCCAATTTGATACCCAGCAACTTTTGA 2518
QY 3972 GTTGTCTCAGATAGTACCTTTATACCAAAACCGTGACGACGAGATTTGGTAGACTCCA 4031
DB 2519 GTTGTCTCAGATAGTACCTTTATACCAAAACCGTGACGACGAGATTTGGTAGACTCCA 2578
QY 4032 GTTGTGTCTTATAGCTTCCGGAATAGACTTTTGGACGAGTACACAGGCCCAACGAG 4091
DB 2579 GTTGTGTCTTATAGCTTCCGGAATAGACTTTTGGACGAGTACACAGGCCCAACGAG 2638
QY 4092 TAATTAGAAGAGTCAGCCACCAAGTAGTGAATAGACCATCGGGGGCTCAGTAGTCAAA 4151

DB 2639 TAATTAGAAGAGTCAGCCACCAAGTAGTGAATAGACCATCGGGGGCTCAGTAGTCAAA 2698
QY 4152 GACGCCAACCAAAATTTTCACTGACAGGGAATTTTTCGACTTCTTTCAGAAAGTTTGGTATTC 4211
DB 2699 GACGCCAACCAAAATTTTCACTGACAGGGAATTTTTCGACTTCTTTCAGAAAGTTTGGTATTC 2758
QY 4212 GTAGTCAATTTCCGAGCATCAATTAATGGGATTTATACAGAAACCAAGTGGAAAGTCA 4271
DB 2759 GTAGTCAATTTCCGAGCATCAATTAATGGGATTTATACAGAAACCAAGTGGAAAGTCA 2818
QY 4272 TCTACCAACTTTCCGGCTCTCAGAAAGGATTAACAGTTCCTACTACGGCATTTAGTGAAA 4331
DB 2819 TCTACCAACTTTCCGGCTCTCAGAAAGGATTAACAGTTCCTACTACGGCATTTAGTGAAA 2878
QY 4332 CTTTTCACCAATTCGCGCTGAGAGAAAGGACACGCGATATCTAGCATTTAGCGGCAAG 4391
DB 2879 CTTTTCACCAATTCGCGCTGAGAGAAAGGACACGCGATATCTAGCATTTAGCGGCAAG 2938
QY 4392 GATGCAACTTTATCAACAGGCTTCTATAGATTAACCTAGCGCTGGGATCATCTTTGG 4451
DB 2939 GATGCAACTTTATCAACAGGCTTCTATAGATTAACCTAGCGCTGGGATCATCTTTGG 2998
QY 4452 ACAACTCTTTCTGCCAAATCTAGGTCACAAATCACITTCATTTGATACCATTTATTGTACAAC 4511
DB 2999 ACAACTCTTTCTGCCAAATCTAGGTCACAAATCACITTCATTTGATACCATTTATTGTACAAC 3058
QY 4512 TTGAGCAAGTTTGTGATCTAGCTCTCTCAATTTGTCTCTGTAAACGATGACTCAACTTGC 4571
DB 3059 TTGAGCAAGTTTGTGATCTAGCTCTCTCAATTTGTCTCTGTAAACGATGACTCAACTTGC 3118
QY 4572 ACATTAACTTGAGAGCTCAGTCGATTCAGTGAATTCATCAGGTTGTGACGTTGTCAGCA 4631
DB 3119 ACATTAACTTGAGAGCTCAGTCGATTCAGTGAATTCATCAGGTTGTGACGTTGTCAGCA 3178
QY 4632 GCATAGGAAACACGGCTTTTCTTACCAAACTCAAGAAATTTATCAAACTCTGCAACACTT 4691
DB 3179 GCATAGGAAACACGGCTTTTCTTACCAAACTCAAGAAATTTATCAAACTCTGCAACACTT 3238
QY 4692 GCATATGACAGTAGCAAGGAAATGTCTATCTTGAAGTCGACAGTGTAGTCTTGA 4751
DB 3239 GCATATGACAGTAGCAAGGAAATGTCTATCTTGAAGTCGACAGTGTAGTCTTGA 3298
QY 4752 GAAATCTCAAGCGCTATTTTATTTATCAGTGTGATCAGTGTGATCCTTACAGCT 4811
DB 3299 GAAATCTCAAGCGCTATTTTATTTATCAGTGTGATCAGTGTGATCCTTACAGCT 3358
QY 4812 GACGATCTGTGGCCGACCTGACAGTTCGCGATCAACGGCGCCACAGTTCGGTGTGCG 4871
DB 3359 GACGATCTGTGGCCGACCTGACAGTTCGCGATCAACGGCGCCACAGTTCGGTGTGCG 3406
QY 4872 GCCTATATCCGACATCAACGATGGGGAAGATCGGCTCGCCACTTCGGGCTCATGAGC 4931
DB 3407 GCCTATATCCGACATCAACGATGGGGAAGATCGGCTCGCCACTTCGGGCTCATGAGC 3466
QY 4932 GCTTGTTCGGGCTGGGTATGGTGGCAGGCCCTTGGCGGGGAGCTGTGGCGGCTATC 4991
DB 3467 GCTTGTTCGGGCTGGGTATGGTGGCAGGCCCTTGGCGGGGAGCTGTGTGGCGGCTATC 3526
QY 4992 TCTTGTGATGACCACTTCTTCTGGCGGGGCTGTCAACGGCTCAACCTACTACTGGGC 5051
DB 3527 TCTTGTGATGACCACTTCTTCTGGCGGGGCTGTCAACGGCTCAACCTACTACTGGGC 3586
QY 5052 TGCTTCTTAAATGACAGGATTCGATAGGAGAGCGTCTAGATATCTATGATTTGAAAGTATG 5111
DB 3587 TGCTTCTTAAATGACAGGATTCGATAGGAGAGCGTCTAGATATCTATGATTTGAAAGTATG 3646
QY 5112 GGAATGGTATACCGCATCTTCTAGTGTCTTGAAGTCTTCTATCAGATTTATGCCCACT 5171
DB 3647 GGAATGGTATACCGCATCTTCTAGTGTCTTGAAGTCTTCTATCAGATTTATGCCCACT 3706
QY 5172 AAAGCAACCGGAGGAGATTTTCAATTTCTGATAAATTTCTCTGACTTTTGGTCACTAGAC 5231

Db 3707 AAAGCAACCGGAGGAGATTTTCATGTAATAATTTCTCTGACTTTTGTGTCATCAGTAGAC 3766
Qy 5232 TCGAACCTGTGAGACTATCTCGGTATATGACAGACAGAAATGTCTCTTTCGGAGACAGTAAT 5291
Db 3767 TCGAACCTGTGAGACTATCTCGGTATATGACAGACAGAAATGTCTCTTTCGGAGACAGTAAT 3826
Qy 5292 GAAGTCCCAACCAATAAAGAAATCTTGTATATCAGGAACAATCTCTCTTTCGAACTTTT 5351
Db 3827 GAAGTCCCAACCAATAAAGAAATCTTGTATATCAGGAACAATCTCTTTCGAACTTTT 3886
Qy 5352 TCGGTGCTTGAACATATAAATATGTAGAGTGNATGTTCGGGTAGGAATGAGACGGGCAAA 5411
Db 3887 TCGGTGCTTGAACATATAAATATGTAGAGTGNATGTTCGGGTAGGAATGAGACGGGCAAA 3946
Qy 5412 TGCTTACCTTCGGACCTTCAAGAGGTATGTAGGGTTTGTAGATACTGATGCCAACTTCA 5471
Db 3947 TGCTTACCTTCGGACCTTCAAGAGGTATGTAGGGTTTGTAGATACTGATGCCAACTTCA 4006
Qy 5472 GTGACAAAGTGTCTATTTTCGTTCAACCAATTCGGAATCCAGAGAAATCAAAAGTTTGTGT 5531
Db 4007 GTGACAAAGTGTCTATTTTCGTTCAACCAATTCGGAATCCAGAGAAATCAAAAGTTTGTGT 4066
Qy 5532 CTACTATTGATCCAGCCAGTGGGTCTTGAACCTGCAATAGTGTGTCTGTGTGTGTGAG 5591
Db 4067 CTACTATTGATCCAGCCAGTGGGTCTTGAACCTGCAATAGTGTGTCTGTGTGTGTGAG 4126
Qy 5592 GTCATCTTTGTATGAATAAATCTAGTCTTGTATCTTGAATATCTTGACGAGCCAAAGCGGA 5651
Db 4127 GTCATCTTTGTATGAATAAATCTAGTCTTGTATCTTGAATATCTTGACGAGCCAAAGCGGA 4186
Qy 5652 TAAATACCCAAATCTTAAATCTTTTAAACCTTAAAGGACAAAGTATGTCTGCTGTAT 5711
Db 4187 TAAATACCCAAATCTTAAATCTTTTAAACCTTAAAGGACAAAGTATGTCTGCTGTAT 4246
Qy 5712 TAAACCCCAATCTGCTGTATCTGATCTCATCAATCTTGAAGGACATATCTTGTGT 5771
Db 4247 TAAACCCCAATCTGCTGTATCTGATCTCATCAATCTTGAAGGACATATCTTGTGT 4306
Qy 5772 AGAGAAATTTGGGAGATCCGATATCGAGAAAGATGACGTGATTTTAAACGTGAAT 5831
Db 4307 AGAGAAATTTGGGAGATCCGATATCGAGAAAGATGACGTGATTTTAAACGTGAAT 4366
Qy 5832 TATCTCAAGATCTGCTGCTGCGGTTCGCTGATGACGGTGAAGAACTCTGACACATGC 5891
Db 4367 TATCTCAAGATCTGCTGCTGCGGTTCGCTGATGACGGTGAAGAACTCTGACACATGC 4426
Qy 5892 AGCTCCGAGACGCTACAGCTTCTGTGAAGCGGATGCGGAGCAGACAAAGCCGCTC 5951
Db 4427 AGCTCCGAGACGCTACAGCTTCTGTGAAGCGGATGCGGAGCAGACAAAGCCGCTC 4486
Qy 5952 AGGGCGCTCAGCGGTGTTCGCGGTTCGCGGCGCAGCCATGACCCAGTCAAGTAGCG 6011
Db 4487 AGGGCGCTCAGCGGTGTTCGCGGTTCGCGGCGCAGCCATGACCCAGTCAAGTAGCG 4546
Qy 6012 ATAGCGAGTGTATCTGCTTAACTATGCGGCATCAGACAGATTTGACTGAGAGTGA 6071
Db 4547 ATAGCGAGTGTATCTGCTTAACTATGCGGCATCAGACAGATTTGACTGAGAGTGA 4606
Qy 6072 CCATATGCGGTGTGAATAACGCAAGATGCGGTGAAGGAGAAATACCGCATCAGGCGCTC 6131
Db 4607 CCATATGCGGTGTGAATAACGCAAGATGCGGTGAAGGAGAAATACCGCATCAGGCGCTC 4666
Qy 6132 TTCGCTTCTCGCTACGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6191
Db 4667 TTCGCTTCTCGCTACGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4726
Qy 6192 AGCTACTCAAAAGGCGGTAAATACGTTATCCACAGAAATCAGGGGATTAACGAGGAAGAA 6251
Db 4727 AGCTACTCAAAAGGCGGTAAATACGTTATCCACAGAAATCAGGGGATTAACGAGGAAGAA 4786
Qy 6252 CATGTGAGCAAAAGGCGCAAAAGGCGCAAGGACCGTAAAGGCGGCTGCTGCTGCTGCTGCT 6311
Db 4787 CATGTGAGCAAAAGGCGCAAAAGGCGCAAGGACCGTAAAGGCGGCTGCTGCTGCTGCTGCT 4846

Qy 6312 TTTTCATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTG 6371
Db 4847 TTTTCATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTG 4906
Qy 6372 GCGAAACCCGACGAGACTATAAAGATACCAAGCGCTTTCCCCCTGGAAGCTCCCTCGTGG 6431
Db 4907 GCGAAACCCGACGAGACTATAAAGATACCAAGCGCTTTCCCCCTGGAAGCTCCCTCGTGG 4966
Qy 6432 CTCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTCGGGAAG 6491
Db 4967 CTCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTCGGGAAG 5026
Qy 6492 CGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTTCGCTGTAGTTCGCTC 6551
Db 5027 CGTGGCGCTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTTCGCTGTAGTTCGCTC 5086
Qy 6552 CHAGCTGGGCTGTGTGACAGAACCCCGCTTCAGCCGACGCGCTGCGCTTATCCGTTAA 6611
Db 5087 CAAGCTGGGCTGTGTGACAGAACCCCGCTTCAGCCGACGCGCTGCGCTTATCCGTTAA 5146
Qy 6612 CTATCTCTTTCAGTCCAAACCCGCTTAAGACACGACTTATCGCCACTGCGACGACCACTGG 6671
Db 5147 CTATCTCTTTCAGTCCAAACCCGCTTAAGACACGACTTATCGCCACTGCGACGACCACTGG 5206
Qy 6672 TAAACAGGATTAAGACGAGGAGGTATGTAGGGGTGTCTACAGAGTTCTTGAAGTGGTGGC 6731
Db 5207 TAAACAGGATTAAGACGAGGAGGTATGTAGGGGTGTCTACAGAGTTCTTGAAGTGGTGGC 5266
Qy 6732 TAACTACCGCTTACACTAAGAGGACAGTATTTGGTATCTGCGCTCTCTGCTAAGCAGTTAC 6791
Db 5267 TAACTACCGCTTACACTAAGAGGACAGTATTTGGTATCTGCGCTCTCTGCTAAGCAGTTAC 5326
Qy 6792 CTTGCGAAAAAGAGTGTGTAGTCTTTCGATCCGGCAACAAACACCGCTGCTGAGCGGTGG 5851
Db 5327 CTTGCGAAAAAGAGTGTGTAGTCTTTCGATCCGGCAACAAACACCGCTGCTGAGCGGTGG 5386
Qy 6852 TTTTCTTTTTCGAAAGGATTTACGCGCAGAAAAAAGGATCTCAAGAGATCCTTT 6911
Db 5387 TTTTCTTTTTCGAAAGGATTTACGCGCAGAAAAAAGGATCTCAAGAGATCCTTT 5446
Qy 6912 GATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACTCAGTTAAGGATTTTGT 5971
Db 5447 GATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACTCAGTTAAGGATTTTGT 5506
Qy 6972 CATGAGATTATCAAAAAGGATCTTACCTAGATCTTTTAAATTTAAATTTAAATTTAA 7031
Db 5507 CATGAGATTATCAAAAAGGATCTTACCTAGATCTTTTAAATTTAAATTTAAATTTAA 5566
Qy 7032 ATCAATCTAAAGTATATAGGTAACTTGTCTGACAGTTACCAATGCTTAATCAGTGA 7091
Db 5567 ATCAATCTAAAGTATATAGGTAACTTGTCTGACAGTTACCAATGCTTAATCAGTGA 5626
Qy 7092 GGCACCTATCTCAGCGATCTGTATTTTCTTATTCATCATAGTTGCTGACTCCCGCTCGT 7151
Db 5627 GGCACCTATCTCAGCGATCTGTATTTTCTTATTCATCATAGTTGCTGACTCCCGCTCGT 5686
Qy 7152 GTAGATACTACGATACGGAGGGCTTACCATCTGCGCCCGAGTGTGCAATGATACCGG 7211
Db 5687 GTAGATACTACGATACGGAGGGCTTACCATCTGCGCCCGAGTGTGCAATGATACCGG 5746
Qy 7212 AGACCCAGCTTACCGGCTCCAGATTATCAGCAATAAACCAGCCAGCGGAGGCGCA 7271
Db 5747 AGACCCAGCTTACCGGCTCCAGATTATCAGCAATAAACCAGCCAGCGGAGGCGCA 5806
Qy 7272 GCGCAGAAAGTGTCTCGAACTTTATCCGCTCCATCCAGTCTTATTTATTTGTTGCGGGA 7331
Db 5807 GCGCAGAAAGTGTCTCGAACTTTATCCGCTCCATCCAGTCTTATTTATTTGTTGCGGGA 5866
Qy 7332 AGCTAGAGTAAAGTGTGCGCAGTTAATAGTTTTCGCAACGTTTGTTCCTATGCTCAGG 7391
Db 5867 AGCTAGAGTAAAGTGTGCGCAGTTAATAGTTTTCGCAACGTTTGTTCCTATGCTCAGG 5926

2927 TAAGGTAACCAATCTTCTTAGTGGAGATGCAAAAACATTTCTTTGCAACCGCACTT 2986
3061 TGGCAGGAACACCCAGCATCAGGAAGTGGAGGAGCAATTTGCGGTTCCACCGGAATAT 3120
2987 TGGCAGGAACACCCAGCATCAGGAAGTGGAGGAGCAATTTGCGGTTCCACCGGAATAT 3046
3121 AGAGCCAACTTTCTCAATAGGTTCTGCAAAACGAGAGCAGACTACACCGGGCAAGTCT 3180
3047 AGAGCCAACTTTCTCAATAGGTTCTGCAAAACGAGAGCAGACTACACCGGGCAAGTCT 3106
3181 CAATCTGCAACAGTCTCCGTTAGTTGAGCTTCATGGAATTTCTGCACTTATCTATAGAGA 3240
3107 CAATCTGCAACAGTCTCCGTTAGTTGAGCTTCATGGAATTTCTGCACTTATCTATAGAGA 3166
3241 GATCAATGGCTCTCTTACAGTTATCTGCAATTTGCAATTTGCTCTGCGGAAGAGGCTT 3300
3167 GATCAATGGCTCTCTTACAGTTATCTGCAATTTGCAATTTGCTCTGCGGAAGAGGCTT 3226
3301 CTAACAGAGTGTCTTCAAGGCACTCCATCAAACTTGGCAGTTAGTTCTTAAAGGGCTT 3360
3227 CTAACAGAGTGTCTTCAAGGCACTCCATCAAACTTGGCAGTTAGTTCTTAAAGGGCTT 3286
3361 TGTCACCAATTTGCAACATTTGTCGACAAATTTGTTGCACTTAATTCATATCTGTTCCG 3420
3287 TGTCACCAATTTGCAACATTTGTCGACAAATTTGTTGCACTTAATTCATATCTGTTCCG 3346
3421 TTTTCTGATAGGAGCAGAGGAGCATCTTCAATTTCTGAGAGGAGGCTTGAAGAGCT 3480
3347 TTTTCTGATAGGAGCAGAGGAGCATCTTCAATTTCTGAGAGGAGGCTTGAAGAGCT 3406
3481 CAATTTTGCACAATTTCAATAGCACTTCAAGAGGACCTCTTTAGTTAGTTGGAATCTTCTT 3540
3407 CAATTTTGCACAATTTCAATAGCACTTCAAGAGGACCTCTTTAGTTAGTTGGAATCTTCTT 3466
3541 TAGTTGTTCTTCTGTTGATCTGCTGATCTGCTGATCTGCTTCTTCTGATGACCTTTAGGG 3600
3467 TAGTTGTTCTTCTGTTGATCTGCTGATCTGCTGATCTGCTTCTTCTGATGACCTTTAGGG 3526
3601 ACTTCATATCCAGGTTCTCTCCACCTGCTCAACGCTCACACGCTGCTGACATCTTAA 3660
3527 ACTTCATATCCAGGTTCTCTCCACCTGCTCAACGCTCACACGCTGCTGACATCTTAA 3586
3661 CTAATGCAAAATAAATAAGTACGACATCTCCAGGCTATATCTTCTTGGATTTAGCTT 3720
3587 CTAATGCAAAATAAATAAGTACGACATCTCCAGGCTATATCTTCTTGGATTTAGCTT 3646
3721 CTGCAAGTTTCATCAGCTTCTCTCCCTAAATTTTAGCGTTCAACAAACTTCGTCAAAATA 3780
3647 CTGCAAGTTTCATCAGCTTCTCTCCCTAAATTTTAGCGTTCAACAAACTTCGTCAAAATA 3706
3781 ACCGTTTGGTATAGAACCTTTGGAGCATTTGCTTTAGCATCCCAAGGTTGCTTCCA 3840
3707 ACCGTTTGGTATAGAACCTTTGGAGCATTTGCTTTAGCATCCCAAGGTTGCTTCCA 3766
3841 TGCTCTAGACCTTTTGGTATAGGCAAGGAGTGGTTCCCAAGTACAGAAACCAA 3900
3767 TGCTCTAGACCTTTTGGTATAGGCAAGGAGTGGTTCCCAAGTACAGAAACCAA 3826
3901 CACTGTTTGTTCACCAACAAATTTCAAGCAGTCTCCATCACAATCCAAATTCGATACCCA 3960
3827 CACTGTTTGTTCACCAACAAATTTCAAGCAGTCTCCATCACAATCCAAATTCGATACCCA 3886
3961 GCAACTTTTGGTATAGTGTGCTGATAGTACCTTTATACCAAAACGTTGACCAAGAT 4020
3887 GCAACTTTTGGTATAGTGTGCTGATAGTACCTTTATACCAAAACGTTGACCAAGAT 3946
4021 GGTAGACTCCAGTTTGTGCTTATAGCTTCCGGAATAGACTTTTGGACGAGTACCA 4080
3947 GGTAGACTCCAGTTTGTGCTTATAGCTTCCGGAATAGACTTTTGGACGAGTACCA 4006
4081 GGCCCAACGAGTAAATAGAGAGTACGCAACCAAGTGTGAATAGACCATCGGGCGGT 4140

4007 GGCCCAACGAGTAATTAGAAGAGTCAAGCCACCAAAAGTAGTGAATAGACCATCGGGCGGT 4066
4141 CAGTAGTCAAAACGCGCAACAAATTTCACTGACAGGGAACCTTTTGTGATCATCTTCAGAAA 4200
4067 CAGTAGTCAAAACGCGCAACAAATTTCACTGACAGGGAACCTTTTGTGATCATCTTCAGAAA 4126
4201 GTTCTGATTTCACTGAGTCAATTTGCGAGCATCAATTAATGGGGATTATACAGAAAGCAACAG 4260
4127 GTTCTGATTTCACTGAGTCAATTTGCGAGCATCAATTAATGGGGATTATACAGAAAGCAACAG 4186
4261 TGGAGTCACTACTACCAACTTTTGGGCTCTCAGAAAAAGCATAAACAGTTCTACTACCGC 4320
4187 TGGAGTCACTACTACCAACTTTTGGGCTCTCAGAAAAAGCATAAACAGTTCTACTACCGC 4246
4321 CATTTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAAAGGCAAGGATCTAGCAT 4380
4247 CATTTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAAAGGCAAGGATCTAGCAT 4306
4381 TAGCGGCAAGGATGCAACTTTTATCAACAGGCTCTATAGATTAACCTTAGCGCTTGGGA 4440
4307 TAGCGGCAAGGATGCAACTTTTATCAACAGGCTCTCTATAGATTAACCTTAGCGCTTGGGA 4366
4441 TCATCTTTGGACAACTCTTTCTGCGCAATCTAGGTCCTAAATCACTTCATTTGATACCAT 4500
4367 TCATCTTTGGACAACTCTTTCTGCGCAATCTAGGTCCTAAATCACTTCATTTGATACCAT 4426
4501 TATTGTACAACTTTGAGCAAGTTTGTGATCAGCTCTCAAAATTTGGTCTCTGTAAACGGATG 4560
4427 TATTGTACAACTTTGAGCAAGTTTGTGATCAGCTCTCTCAAAATTTGGTCTCTGTAAACGGATG 4486
4561 ACTCAACTTTGACATTAACCTTTGAGGCTCAGTTCGATTAAGTGAACCTTTGATCAGGTTGTGCA 4620
4487 ACTCAACTTTGACATTAACCTTTGAGGCTCAGTTCGATTAAGTGAACCTTTGATCAGGTTGTGCA 4546
4621 GCTGGTCAAGCATAGGGAACACCGGCTTTTCTCAACAACTCAAGGAATTTATCAAACT 4680
4547 GCTGGTCAAGCATAGGGAACACCGGCTTTTCTCAACAACTCAAGGAATTTATCAAACT 4606
4681 CTGCAACACTTTGATGAGTACAGGAGGAAATGTCATCTTGAAGTGAAGTCAAGGATGAG 4740
4607 CTGCAACACTTTGATGAGTACAGGAGGAAATGTCATCTTGAAGTGAAGTCAAGGATGAG 4666
4741 TGAGTCTTGAAGAAATTTCTGAAGCGGTATTTTATTATCAGTGAAGTCAAGGATGAG 4800
4667 TGAGTCTTGAAGAAATTTCTGAAGCGGTATTTTATTATCAGTGAAGTCAAGGATGAG 4726
4801 TCCTTACGCGGAGCATCGTGGCGGA----- 4828
4727 TCCTTACGCGGAGCATCGTGGCGGAACCTGCAAGGGGGGGGGGGGGGCTGAGGTTGCG 4786
4829 ----- 4828
4787 CTGCTGAAGAGGTTGTGCTGATCATACCAGGCTTGAATCGCCCATCATCCAGCCAGA 4846
4829 ----- 4828
4847 AAGTGGAGGAGCCAGGTTGATGAGAGCTTTGTTGTAGTGGACCAAGTTGGTGTATTTGA 4906
4829 ----- 4828
4907 ACTTTTGTGTTGCCACGGAACGGTCTGCTGTTGTCGGGAAGATGGTGTATCTGATCTTCA 4966
4829 ----- 4828
4967 ACTCAGCAAAAGTTTCGATTTTATTCAACAAAGCCCGCTGCGGTCAAGTCAAGCGTAAATGCT 5026
4829 ----- 4828
5027 CTGCGAGTGTTCACCAACCAATTAACCAATCTGATTAGAAAAAATCATCTCGAGCATCAAAATG 5086
4829 ----- 4828
5087 AAATGCAATTTATTATATATCAGGATTAATCAATACCATATTTTGTAAAAAGCCGTTTCTG 5146

QY 4829 ----- 4828
Db 5147 TAATGAAGGAGAAAACTCACCGAGGCGAGTTCCATAGGATGCAAGATCCTGGTATCGGTC 5206
QY 4829 ----- 4828
Db 5207 TCGGATCCGACTCGTCAACATCAATCAACCTATTAAATTCCTCCCTCGTCAAAAAATAAG 5266
QY 4829 ----- 4828
Db 5267 GTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAAATGCAAAAGCTT 5326
QY 4829 ----- 4828
Db 5327 ATGCATTTCTTTCCAGACTTGTTCACAGGCCAGCCATTACGCTCGTCAATCAAAAATCACT 5386
QY 4829 ----- 4828
Db 5387 CGCATCAACCAACCGTTATTTCATTCTGTGATTGCGCTGAGCGAGACGAATACCGGATC 5446
QY 4829 ----- 4828
Db 5447 GCTGTAAAGGACAAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAG 5506
QY 4829 ----- 4828
Db 5507 CGCATCAACAAATATTTTCACTGTAATCAGGATATTCTTCTAATACCTGGAAATGCTGTTTT 5566
QY 4829 ----- 4828
Db 5567 CCGGGGATCGAGTGTGAGTAAACCATGCATCATCAGGAGTACGGATATAAATGCTTGAT 5626
QY 4829 ----- 4828
Db 5627 GGTGGAGAGGCAATAAATTCGTGAGCCAGTTTATGTCGACCATCATCTGTGTAACATC 5686
QY 4829 ----- 4828
Db 5687 ATTGCAACGCTACCTTTGCCATGTTTCAGAAACAACCTCTGGCGCATCGGGCTTCCCAT 5746
QY 4829 ----- 4828
Db 5747 CAATCGATAGATGTGCGACCTGATTGCGCGACATTATCGCGGCCCATTTATACCCATA 5806
QY 4829 ----- 4828
Db 5807 TAAATCAGCATCCATGTTGGAAATTAATTCGGCGCTCGAGCAAGACGTTTCCCGTTGAAT 5866
QY 4829 ----- 4828
Db 5867 ATGGCTCAACACCCCTTGTTACTGTTTATGTAAGCAGACAGTTTATTGTTTCATGA 5926
QY 4829 ----- 4828
Db 5927 TGATATATTTTATCTGTGCAATGTAAATCATCAGAGATTTTGAGACACAAGTGGCTTTC 5986
QY 4829 ----- CTTGCAAGTTCGCATCACCGCGCCACAGGTGCGGTGCTGCGCGCTATAT 4879
Db 5987 CCCCCCCCCCTGCAAGTTCGGCATCACCGCGCCACAGGTGCGGTGCTGCGCGCTATAT 6046
QY 4880 CGCGACATCACCGATGGGAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTGTTT 4939
Db 6047 CGCGACATCACCGATGGGAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTGTTT 6106
QY 4940 CGGGCTGGGTATGTGGCAGCCCGCTGGCGCGGGGACTGTTGGCGCCCATCTCTTGCA 4999
Db 6107 CGGGCTGGGTATGTGGCAGCCCGCTGGCGCGGGGACTGTTGGCGCCCATCTCTTGCA 6166
QY 5000 TGCACCATTCCTTGGCGGGCGGTGCTCAACGGGCTCAACCTACTACTGGGCTGCTTCCT 5059
Db 6167 TGCACCATTCCTTGGCGGGCGGTGCTCAACGGGCTCAACCTACTACTGGGCTGCTTCCT 6226

QY 5060 AATGCAGAGTCGCATTAAGGAGAGCGTGCAGTATCTATGATTGGAAGTATGGGAATGGT 5119
Db 6227 AATGCAGAGTCGCATTAAGGAGAGCGTGCAGTATCTATGATTGGAAGTATGGGAATGGT 6286
QY 5120 GATACCCGATTTCTTCAGTGTCTTCAGGTCTCCTATCAGATTATCCCAACTTAAAGCAAC 5179
Db 6287 GATACCCGATTTCTTCAGTGTCTTCAGGTCTCCTATCAGATTATCCCAACTTAAAGCAAC 6346
QY 5180 CGAGGAGAGATTTCTATGGTAAATTTCTTGACTTTTGGTCAATCAGTAGACTCGAATG 5239
Db 6347 CGAGGAGAGATTTCTATGGTAAATTTCTTGACTTTTGGTCAATCAGTAGACTCGAATG 6406
QY 5240 TGAGACTATCTCGGTATGACAGCAGAAATGCTTCTTGAGAGCAGTAAATGAAGTCCC 5299
Db 6407 TGAGACTATCTCGGTATGACAGCAGAAATGCTTCTTGAGAGCAGTAAATGAAGTCCC 6466
QY 5300 ACCAATAAAGAAATCCTTGTATCAGGAACAAACTTCTTGTTCGAACTTTTTCGGTCCC 5359
Db 6467 ACCAATAAAGAAATCCTTGTATCAGGAACAAACTTCTTGTTCGAACTTTTTCGGTCCC 6526
QY 5360 TTGAATATAAATGATAGTGGATATGTCGGTAGGAATGCGCGGCAATGCTTACC 5419
Db 6527 TTGAATATAAATGATAGTGGATATGTCGGTAGGAATGCGCGGCAATGCTTACC 6586
QY 5420 TTCTGGACCTTCAAGAGGTATGAGGTTTGTAGATACTGATGCCAACTTCAGTGACAAC 5479
Db 6587 TTCTGGACCTTCAAGAGGTATGAGGTTTGTAGATACTGATGCCAACTTCAGTGACAAC 6646
QY 5480 GTTGCTATTTCTTCAAAACCATTCGAATTCAGAGAAATCAAAAGTTGTTGTCTACTATT 5539
Db 6647 GTTGCTATTTCTTCAAAACCATTCGAATTCAGAGAAATCAAAAGTTGTTGTCTACTATT 6706
QY 5540 GATCAAGCAGTGGCTTCTGAAACTGACAATAGTGTCTGCTGTTTGTAGGTCATCTT 5599
Db 6707 GATCAAGCAGTGGCTTCTGAAACTGACAATAGTGTGCTGCTGTTTGTAGGTCATCTT 6766
QY 5600 TGTATGAATAAATCTAGTCTTTGATCTATAATTAATCTTCAAGAGCAAGCGCAATAATACC 5659
Db 6767 TGTATGAATAAATCTAGTCTTTGATCTATAATTAATCTTCAAGAGCAAGCGCAATAATACC 6826
QY 5660 CAATCTAATAACTCTTTTAAACGTTTAAAGCAGTAAGTATGTCGCTGTATTTAAACCC 5719
Db 6827 CAATCTAATAACTCTTTTAAACGTTTAAAGCAGTAAGTATGTCGCTGTATTTAAACCC 6886
QY 5720 AAATCAGCTCGTAGTCTGATCCTCATCAACTTGAAGGGCACTATCTTGTTTTAGAGAAAT 5779
Db 6887 AAATCAGCTCGTAGTCTGATCCTCATCAACTTGAAGGGCACTATCTTGTTTTAGAGAAAT 6946
QY 5780 TTGCGGAGTGGGATATCGAGAAAGGTACGCTGATTTTAAACGTGAATTTATCTCAA 5839
Db 6947 TTGCGGAGTGGGATATCGAGAAAGGTACGCTGATTTTAAACGTGAATTTATCTCAA 7006
QY 5840 GATCTCTGCTCGCGCTTTCGGTGTATGACCGTGAACCTCTGACACATGACGCTCCCG 5899
Db 7007 GATCTCTGCTCGCGCTTTCGGTGTATGACCGTGAACCTCTGACACATGACGCTCCCG 7066
QY 5900 GAGACGGTCAAGCTTGTCTGTAAGCGGATGCGGGAGCAGACAGCCCTCAGGCGCG 5959
Db 7067 GAGACGGTCAAGCTTGTCTGTAAGCGGATGCGGGAGCAGACAGCCCTCAGGCGCG 7126
QY 5960 TCAGCGGTTTGGCGGTGTCGGGCGCAGCATGACCCAGTACGTAGCGATAGCGGA 6019
Db 7127 TCAGCGGTTTGGCGGTGTCGGGCGCAGCCATGACCCAGTACGTAGCGATAGCGGA 7186
QY 6020 GTGTACTGCTTAACATATCGGCAATCAGAGCAGATTGTACTGAGAGTGCACCATATGC 6079
Db 7187 GTGTACTGCTTAACATATCGGCAATCAGAGCAGATTGTACTGAGAGTGCACCATATGC 7246
QY 6080 GGTGTGAATACCGCAGATCGGTAAAGGAAATACCGCATCAGGCGCTTCTCCGTT 6139
Db 7247 GGTGTGAATACCGCAGATCGGTAAAGGAAATACCGCATCAGGCGCTTCTCCGTT 7306
QY 6140 CCTCGCTACTGACTGCTCGTGGCTCGGTGCTTGGCTGCGGCGAGCGGTATCAGCTCACT 6199

LOCUS AX149597 9276 bp DNA linear PAT 31-AUG-2001
DEFINITION Sequence 21 from Patent WO0136604.
ACCESSION AX149597
VERSION AX149597.1 GI:14348030
KEYWORDS Pichia pastoris
SOURCE Pichia pastoris
ORGANISM Pichia pastoris
REFERENCE 1
AUTHORS Madison, E.L. and Ong, E.O.
TITLE Nucleic acids encoding endotheliases, endotheliases and uses
JOURNAL Patent: WO 0136604-A 21 25-MAY-2001;
CORVAS INTERNATIONAL, INC. (US)
FEATURES
source Location/Qualifiers
1..9276
/organism="Pichia pastoris"
/mol_type="unassigned DNA"
/db_xref="taxon:4922"
/note="pPic9K vector"
ORIGIN
Query Match 79.6%; Score 6561.6; DB 6; Length 9276;
Best local Similarity 84.7%; Pred. No. 0;
Matches 8031; Conservative 0; Mismatches 4; Indels 1447; Gaps 2;
QY 1 AGATCTACATCCAAAGACGAAAGGTTGAATGAACCTTTTGGCATCCGACATCCACAG 60
DB 1 AGATCTACATCCAAAGACGAAAGGTTGAATGAACCTTTTGGCATCCGACATCCACAG 60
QY 61 GTCCATTCTCACATAAGTCCCAAAACGCAACAGAGGGGATACATAGCAGACAGCGT 120
DB 61 GTCCATTCTCACATAAGTCCCAAAACGCAACAGAGGGGATACATAGCAGACAGCGT 120
QY 121 TGCAAACGAGACCTCAGCT 180
DB 121 TGCAAACGAGACCTCAGCT 180
QY 181 AGCCAGATTATGGCTTGATGGAGTCTGCTCATTCCTCTCTCTCTCTCTCTCTCTCTCT 240
DB 181 AGCCAGATTATGGCTTGATGGAGTCTGCTCATTCCTCTCTCTCTCTCTCTCTCTCTCT 240
QY 241 ACACCATGACTTTATAGCTTGTCTATCTGGCCCCCTGCGAGGTTCTATGTTGTTTA 300
DB 241 ACACCATGACTTTATAGCTTGTCTATCTGGCCCCCTGCGAGGTTCTATGTTGTTTA 300
QY 301 TTTCGGAATGCAACAGCTCCGCAATTACCCGGAACATCTCCAGATGAGGGCTTTCTG 360
DB 301 TTTCGGAATGCAACAGCTCCGCAATTACCCGGAACATCTCCAGATGAGGGCTTTCTG 360
QY 361 AGTGTGGGGTCAAAATAGTTTCATGTTCCCAAAATGGCCCAAAATCTGACAGTTTAAACGCT 420
DB 361 AGTGTGGGGTCAAAATAGTTTCATGTTCCCAAAATGGCCCAAAATCTGACAGTTTAAACGCT 420
QY 421 GTCTTGGAACTTAATGCAAAAGCGTGATCTCATCCAGATGAATAGTTGGTTCG 480
DB 421 GTCTTGGAACTTAATGCAAAAGCGTGATCTCATCCAGATGAATAGTTGGTTCG 480
QY 481 TTGAAATGCTAACGCCAGTTGGTCAAAAGAAATCTCCAAAGTCCGCCATACCGTTTGT 540
DB 481 TTGAAATGCTAACGCCAGTTGGTCAAAAGAAATCTCCAAAGTCCGCCATACCGTTTGT 540
QY 541 CTGTGTTGGTATGATGAGCAATGCTCAAAATAATCTATTAATGCTTAGCGCAGTCT 600
DB 541 CTGTGTTGGTATGATGAGCAATGCTCAAAATAATCTATTAATGCTTAGCGCAGTCT 600
QY 601 CTCATATCGCTTCTGAACCCCGGTGCACTGTGCGGAAACGCAATGGGGAACACCCGCT 660
DB 601 CTCATATCGCTTCTGAACCCCGGTGCACTGTGCGGAAACGCAATGGGGAACACCCGCT 660
QY 661 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 720
DB 661 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 720
QY 721 GCTGATAGCTTAACCTTCATGATCAAAATTTAACTGTCTAAACCTTACTTACGAGCAAT 780
DB 721 GCTGATAGCTTAACCTTCATGATCAAAATTTAACTGTCTAAACCTTACTTACGAGCAAT 780
QY 781 ATATAACAGAGGAGGAGCTGCCCTGTCTTAAACCTTTTATCATCATATTAGCTT 840
DB 781 ATATAACAGAGGAGGAGCTGCCCTGTCTTAAACCTTTTATCATCATATTAGCTT 840
QY 841 ACTTTCATAATTCGACCTGGTTTCAAAATGCAAGCTTTTGAATTTAAGCACTTTAAGCA 900
DB 841 ACTTTCATAATTCGACCTGGTTTCAAAATGCAAGCTTTTGAATTTAAGCACTTTAAGCA 900
QY 901 CAACTTGAGAGATCAAAACCACTAATTTTGAAGGATCCAAACGATGAGATTTCCT 960
DB 901 CAACTTGAGAGATCAAAACCACTAATTTTGAAGGATCCAAACGATGAGATTTCCT 960
QY 961 TCAATTTTACTGCACTTTTATTCGACATCTCCGCAATTCCTCCAGTCAACACT 1020
DB 961 TCAATTTTACTGCACTTTTATTCGACATCTCCGCAATTCCTCCAGTCAACACT 1020
QY 1021 ACAACAGAGATGAACGGCACAAATTCGGCTGAGCTGTCATCGGTTACTCAGATTTA 1080
DB 1021 ACAACAGAGATGAACGGCACAAATTCGGCTGAGCTGTCATCGGTTACTCAGATTTA 1080
QY 1081 GAAAGGGGATTCGATGTTGCTGTTTGGCCATTTTCCAAACAGCACAAATAACGGGTTATTG 1140
DB 1081 GAAAGGGGATTCGATGTTGCTGTTTGGCCATTTTCCAAACAGCACAAATAACGGGTTATTG 1140
QY 1141 TTTTAAATATCTATTTGCGCAGATTCCTGCTTAAAGAGGGGTATCTCTCGAGAA 1200
DB 1141 TTTTAAATATCTATTTGCGCAGATTCCTGCTTAAAGAGGGGTATCTCTCGAGAA 1200
QY 1201 ACAGAGGCTGAAGCCAGGCCAGGCCAGGCCAGGCCAGGAAACATGATCATG 1260
DB 1201 ACAGAGGCTGAAGCCAGGCCAGGCCAGGCCAGGCCAGGAAACATGATCATG 1260
QY 1261 GCCCCCCGGGAGGAGATAAATTTGTGGCTTCCCGGTGTCACCGCCAGCAGTGCAGGAG 1320
DB 1261 GCCCCCCGGGAGGAGATAAATTTGTGGCTTCCCGGTGTCACCGCCAGCAGTGCAGGAG 1320
QY 1321 AGAGGTTGCTGTTTGTATGACAGTCTCCGGGATTCCTGCTTCCACCCCATGSCC 1380
DB 1321 AGAGGTTGCTGTTTGTATGACAGTCTCCGGGATTCCTGCTTCCACCCCATGSCC 1380
QY 1381 ATCGAGAACTCAAGAAAGAAATGTCCTTCTAACTAGTGGCGTAGAATTCCTAGGG 1440
DB 1381 ATCGAGAACTCAAGAAAGAAATGTCCTTCTAACTAGTGGCGTAGAATTCCTAGGG 1440
QY 1441 CGGCGCGAATTAATTCGCTTAGCATGACTGTTCTCTCAGTTCAAGTTGGGCACTTACG 1500
DB 1441 CGGCGCGAATTAATTCGCTTAGCATGACTGTTCTCTCAGTTCAAGTTGGGCACTTACG 1500
QY 1501 AGAAGACCGCTCTGCTAGATTCTAATCAAGAGGATGTCAGAAATGCCATTTGCGCTGAGAG 1560
DB 1501 AGAAGACCGCTCTGCTAGATTCTAATCAAGAGGATGTCAGAAATGCCATTTGCGCTGAGAG 1560
QY 1561 ATGCGGCTTCATTTTGTATCTTTTATTTTGTAACTATATAGTATAGGATTTTCTT 1620
DB 1561 ATGCGGCTTCATTTTGTATCTTTTATTTTGTAACTATATAGTATAGGATTTTCTT 1620
QY 1621 GTCAATTTGTTTCTTCTCGTAGAGCTTGCCTGATCAGCCTATCTCGAGCTGATGAA 1680
DB 1621 GTCAATTTGTTTCTTCTCGTAGAGCTTGCCTGATCAGCCTATCTCGAGCTGATGAA 1680
QY 1681 TATCTTGTAGGAGTTTGGGAAATCATTCGAGTTTGTATGTTTCTTCTGTTATTTCCC 1740
DB 1681 TATCTTGTAGGAGTTTGGGAAATCATTCGAGTTTGTATGTTTCTTCTGTTATTTCCC 1740
QY 1741 ACTCTCTCTTCAGAGTACAGAGATTAAGTGAGAGTTTCTGTTTGTGCAAGCTTATCGATAA 1800
DB 1741 ACTCTCTCTTCAGAGTACAGAGATTAAGTGAGAGTTTCTGTTTGTGCAAGCTTATCGATAA 1800
QY 1801 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 1860
DB 1801 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 1860
QY 1861 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 1920
DB 1861 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 1920
QY 1921 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 1980
DB 1921 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 1980
QY 1981 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2040
DB 1981 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2040
QY 2041 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2100
DB 2041 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2100
QY 2101 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2160
DB 2101 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2160
QY 2161 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2220
DB 2161 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2220
QY 2221 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2280
DB 2221 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2280
QY 2281 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2340
DB 2281 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2340
QY 2341 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2400
DB 2341 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2400
QY 2401 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2460
DB 2401 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2460
QY 2461 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2520
DB 2461 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2520
QY 2521 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2580
DB 2521 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2580
QY 2581 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2640
DB 2581 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2640
QY 2641 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2700
DB 2641 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2700
QY 2701 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2760
DB 2701 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2760
QY 2761 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2820
DB 2761 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2820
QY 2821 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2880
DB 2821 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2880
QY 2881 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2940
DB 2881 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 2940
QY 2941 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3000
DB 2941 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3000
QY 3001 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3060
DB 3001 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3060
QY 3061 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3120
DB 3061 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3120
QY 3121 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3180
DB 3121 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3180
QY 3181 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3240
DB 3181 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3240
QY 3241 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3300
DB 3241 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3300
QY 3301 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3360
DB 3301 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3360
QY 3361 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3420
DB 3361 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3420
QY 3421 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3480
DB 3421 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3480
QY 3481 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3540
DB 3481 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3540
QY 3541 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3600
DB 3541 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3600
QY 3601 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3660
DB 3601 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3660
QY 3661 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3720
DB 3661 TTTTGGATGATTATGCAATGCTCCAAATTTGATGCTTCCAAATTTCTGGTGGGAATACT 3720

QY	1801	GC	TTAATGCGGTAGTTTATCAGAGTTTAAATTGTCTAACGCAGTCAGCGACCGGTGTATGAA	1860
DB	1595	GC	TTAATGCGGTAGTTTATCAGAGTTTAAATTGTCTAACGCAGTCAGCGACCGGTGTATGAA	1654
QY	1861	AT	CTAACAAATGCGGTCTATCGTCATCTCTCGGCAACCGTCACCCCTGGATGCTGTAGGCATPAGG	1920
DB	1655	AT	CTAACAAATGCGGTCTATCGTCATCTCTCGGCAACCGTCACCCCTGGATGCTGTAGGCATPAGG	1714
QY	1921	CT	TGGTTATGCGGTATCTGCGGGGCCCTTTGGGGATATCGTCCATTCCGACAGCATCGC	1980
DB	1715	CT	TGGTTATGCGGTATCTGCGGGGCCCTTTGGGGATATCGTCCATTCCGACAGCATCGC	1774
QY	1981	CAG	TCATCTAGCGGTGCTGTCTAGCGCTATATGGGTTGATGCAATTTCTATGCGCAACCCGT	2040
DB	1775	CAG	TCATCTAGCGGTGCTGTCTAGCGCTATATGGGTTGATGCAATTTCTATGCGCAACCCGT	1834
QY	2041	TC	TCGGAGCACTGTCGACACCGCTTTGGCGCGCGCCAGTCTCGTCCGCTTCGCTACTCTGG	2100
DB	1835	TC	TCGGAGCACTGTCGACACCGCTTTGGCGCGCGCCAGTCTCGTCCGCTTCGCTACTCTGG	1894
QY	2101	AG	CCACTATCGACTACGCGATCATGGCGACACACACCCGTCCTGTGGATCTATCGAATCTTA	2160
DB	1995	AG	CCACTATCGACTACGCGATCATGGCGACACACACCCGTCCTGTGGATCTATCGAATCTTA	1954
QY	2161	AA	TGTAAAGTTAAATCTCTAAATTAATTAATTAAGTCCAGTTTCTCAATAGCAACTTAA	2220
DB	1955	AA	TGTAAAGTTAAATCTCTAAATTAATTAATTAAGTCCAGTTTCTCAATAGCAACTTAA	2014
QY	2221	CAG	CAATTCGGGTAGCATCTAGACCTTTCAACAGCAGCGACATCCATCACTGCTTGGCCAA	2280
DB	2015	CAG	CAATTCGGGTAGCATCTAGACCTTTCAACAGCAGCGACATCCATCACTGCTTGGCCAA	2074
QY	2281	TA	TTTTTCAGTCCCTCAGGAGTTACGTCCTGTGAAAGTGATGAATCTCTGGAAGTTCCAG	2340
DB	2075	TA	TTTTTCAGTCCCTCAGGAGTTACGTCCTGTGAAAGTGATGAATCTCTGGAAGTTCCAG	2134
QY	2341	TG	TTAACTCCGCTGATTTGACGGGCATATCGTAGCTTGGCAAGTGTGGTTGTGTACCGG	2400
DB	2135	TG	TTAACTCCGCTGATTTGACGGGCATATCGTAGCTTGGCAAGTGTGGTTGTGTACCGG	2194
QY	2401	AG	GAGTAATCTCCAACACTCTCTCGAGAGTAGGCCAACCAACACACAGATCCAGCGTGT	2460
DB	2195	AG	GAGTAATCTCCAACACTCTCTCGAGAGTAGGCCAACCAACACACAGATCCAGCGTGT	2254
QY	2461	GT	ACTTGATCAACATAGACAGCAATCTCGATTTGCAGGATCAAGTGTTCAGGACGT	2520
DB	2255	GT	ACTTGATCAACATAGACAGCAATCTCGATTTGCAGGATCAAGTGTTCAGGACGT	2314
QY	2521	ACT	GATTTGGCAATTTCCAAGCCCTGCTAGGTTGCAACCGATAGGGTTGTAGAGTGTG	2580
DB	2315	ACT	GATTTGGCAATTTCCAAGCCCTGCTAGGTTGCAACCGATAGGGTTGTAGAGTGTG	2374
QY	2581	CA	ATACACTTCGGTACAAATTTCAACCCCTTGGCAACTGCAACGCTTGGTTGTGAACAGCAT	2640
DB	2375	CA	ATACACTTCGGTACAAATTTCAACCCCTTGGCAACTGCAACGCTTGGTTGTGAACAGCAT	2434
QY	2641	CT	TCAAATTCGGCAAGCTCTGCTGTGTCATATCGACAGCCCAACAGAAATCACCTGGCAAT	2700
DB	2435	CT	TCAAATTCGGCAAGCTCTGCTGTGTCATATCGACAGCCCAACAGAAATCACCTGGCAAT	2494
QY	2701	CA	ATACCAATGTCAGCTTGACACAGAAAGGTCTGAGGCAACGAAATCTGGATCAGCGTATT	2760
DB	2495	CA	ATACCAATGTCAGCTTGACACAGAAAGGTCTGAGGCAACGAAATCTGGATCAGCGTATT	2554
QY	2761	TAT	CAGCAATTAACATAGACTTTCAGAGGCCAGCAGGATCTCAATACTACACAGGGCTG	2820
DB	2555	TAT	CAGCAATTAACATAGACTTTCAGAGGCCAGCAGGATCTCAATACTACACAGGGCTG	2614
QY	2821	AT	GTGTCAATTTGAAACCATCATCTTTGGCAGCAGTAACGAACTGGTTCTTGGACCAATA	2880
DB	2615	AT	GTGTCAATTTGAAACCATCATCTTTGGCAGCAGTAACGAACTGGTTCTTGGACCAATA	2674

QY	2881	TTTTTGTACACTTTAGGAACAGTTTCTTGTTCCTGTAAGCCATAGCAGCTACTGCTGTGGCGC	2941
DB	2675	TTTTTGTACACTTTAGGAACAGTTTCTTGTTCCTGTAAGCCATAGCAGCTACTGCTGTGGCGC	2734
QY	2841	CTCCTGTAGCAGATACACTTATAGCAACCACTTGTGGCAACGTTAGACTTCTCTGGG	3000
DB	2735	CTCCTGTAGCAGATACACTTATAGCAACCACTTGTGGCAACGTTAGACTTCTCTGGG	2794
QY	3001	TAAGGGTACCATTCTTCTTAGTGGAGATGCAAAAACAATTTCTTTTGCAACCGAACACTT	3060
DB	2795	TAAGGGTACCATTCTTCTTAGTGGAGATGCAAAAACAATTTCTTTTGCAACCGAACACTT	2854
QY	3061	TGGCAGAACACCCAGCATCAGGGAAGTGAAAGGCAGAAATTTGGGTTCACACAGGAATAT	3120
DB	2855	TGGCAGAAACCCAGCATCAGGGAAGTGAAAGGCAGAAATTTGGGTTCACACAGGAATAT	2914
QY	3121	AGAGGCCAACTTCTCAATAGGCTTTCGAAAAAGAGACGAGATACACACAGGCGCAAGTCT	3180
DB	2915	AGAGGCCAACTTCTCAATAGGCTTTCGAAAAAGAGACGAGATACACACAGGCGCAAGTCT	2974
QY	3181	CAACTTGCACAGTCTCGTTAGTTGAGCTTCAATGGAAATTTCTTGAAGTTCCTATAGAGA	3240
DB	2975	CAACTTGCACAGTCTCGTTAGTTGAGCTTCAATGGAAATTTCTTGAAGTTCCTATAGAGA	3034
QY	3241	GATCAATGGCTCTCTTAACGTTATCTCGCAATTTGCATAAGTTCTCTGTGGAAAGGAGCTT	3300
DB	3035	GATCAATGGCTCTCTTAACGTTATCTCGCAATTTGCATAAGTTCTCTGTGGAAAGGAGCTT	3094
QY	3301	CTAACACAGGTGTCTTCAAGCGACTCCATCAAACTTTGGCAGTTAGTTCTTAAAGGGCTT	3360
DB	3095	CTAACACAGGTGTCTTCAAGCGACTCCATCAAACTTTGGCAGTTAGTTCTTAAAGGGCTT	3154
QY	3361	TGTCACCAATTTGACGAACATTTGTCGAATTTGGTTTTCATTTCCATATCTGTTCGG	3420
DB	3155	TGTCACCAATTTGACGAACATTTGTCGAATTTGGTTTTCATTTCCATATCTGTTCGG	3214
QY	3421	TTTTTCTGGATAGGACGACGAAGGCACTTCAATTTCTTGTGAGGAGCCCTTAGAAACGT	3480
DB	3215	TTTTTCTGGATAGGACGACGAAGGCACTTCAATTTCTTGTGAGGAGCCCTTAGAAACGT	3274
QY	3481	CAATTTTGCACAAATTCATACGACCTTCAGAGGCACTTCTTTAGGTTTGGATCTCTT	3540
DB	3275	CAATTTTGCACAAATTCATACGACCTTCAGAGGCACTTCTTTAGGTTTGGATCTCTT	3334
QY	3541	TAGGTTGTTCTCGTGTATCTCGCTTGGCATCTCCTTTTCTTCTAGTGACCTTTTAGGG	3600
DB	3335	TAGGTTGTTCTCGTGTATCTCGCTTGGCATCTCCTTTTCTTCTAGTGACCTTTTAGGG	3394
QY	3601	ACTTCAATATCCAGGTTTCTCTCCACCTTGTCTCCAAAGTCAACCGTACTTTGGCACTCTAA	3660
DB	3395	ACTTCAATATCCAGGTTTCTCTCCACCTTGTCTCCAAAGTCAACCGTACTTTGGCACTCTAA	3454
QY	3661	CTAATGCAAAATAAATAAGTCACACATTTCCAGGCTATATCTTCTTGGATTTAGCTT	3720
DB	3455	CTAATGCAAAATAAATAAGTCACACATTTCCAGGCTATATCTTCTTGGATTTAGCTT	3514
QY	3721	CTGCAAGTTTCATCAGCTTCTCCCTTAATTTTAGCGTTTCAACAAAACCTTCGTCGCAATA	3780
DB	3515	CTGCAAGTTTCATCAGCTTCTCCCTTAATTTTAGCGTTTCAACAAAACCTTCGTCGCAATA	3574
QY	3781	ACCGTTTGGTATAGAACCTTCTGGAGCATGCTCTTAGCATCCCAAGGTGGCTTCCA	3840
DB	3575	ACCGTTTGGTATAGAACCTTCTGGAGCATGCTCTTAGCATCCCAAGGTGGCTTCCA	3634
QY	3841	TGGCTCTAAGACCCCTTTGATTTGGGCCAAAAAGGAAGTGCCTTCCAAGTGAAGAAAAACAA	3900
DB	3635	TGGCTCTAAGACCCCTTTGATTTGGGCCAAAAAGGAAGTGCCTTCCAAGTGAAGAAAAACAA	3694
QY	3901	CACCTGTTTGTTCACACAAATTTCAAGCAGTCTCCATCAATATCCATTCGATGCCA	3960
DB	3695	CACCTGTTTGTTCACACAAATTTCAAGCAGTCTCCATCAATATCCATTCGATGCCA	3754
QY	3961	GCACCTTTTGTAGTTTGTCTCCAGATGTAGCACCTTTTATACCAAAACCGTGACGACGATTT	4020

Db 3755 GCAACTTTTGAGTTGCTCCAGATGTAGCACTTTTATACCAAAACCGTGACGAGATT 3814
Qy 4021 GGTAGACTCCAGTTTGTGCTTATAGCTCCGGATAGACTTTTGGAGGAGTACACCA 4080
Db 3815 GGTAGACTCCAGTTTGTGCTTATAGCTCCGGATAGACTTTTGGAGGAGTACACCA 3874
Qy 4081 GGCCCAAGAGTAATTAGAGAGTACGACCAACCAAGTAGTATAGACCAATCGGGGGGT 4140
Db 3875 GGCCCAAGAGTAATTAGAGAGTACGACCAACCAAGTAGTATAGACCAATCGGGGGGT 3934
Qy 4141 CAGTAGTCAAAAGAGCCCAAAATTTTCACTGACAGGGAATTTTTCACATCTTCAGAAA 4200
Db 3935 CAGTAGTCAAAAGAGCCCAAAATTTTCACTGACAGGGAATTTTTCACATCTTCAGAAA 3994
Qy 4201 GTTCGTATTTCAGTAGTCAATTTGCCAGCATCAATTAATGGGATPATACCAAGAACAG 4260
Db 3995 GTTCGTATTTCAGTAGTCAATTTGCCAGCATCAATTAATGGGATPATACCAAGAACAG 4054
Qy 4261 TGGAGTCACATCTACCAACTTTGGGTCTCAGAAAAGCATAAACAGTTTCTACTCCG 4320
Db 4055 TGGAGTCACATCTACCAACTTTGGGTCTCAGAAAAGCATAAACAGTTTCTACTCCG 4114
Qy 4321 CATTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAAGGCACAGCGATACTAGCAT 4380
Db 4115 CATTAGTGAACCTTTTCAAAATCGCCAGTGGAGAGAAAAGGCACAGCGATACTAGCAT 4174
Qy 4381 TAGCGGGAAGATGCAACTTTTATCAACAGGGTCTTATAGATTAACCTAGCGCTGGGA 4440
Db 4175 TAGCGGGAAGATGCAACTTTTATCAACAGGGTCTTATAGATTAACCTAGCGCTGGGA 4234
Qy 4441 TCATCCTTTGGACAACTTTTCTGCCAAATCTAGTCCAAATCACITTCATTGATACCAT 4500
Db 4235 TCATCCTTTGGACAACTTTTCTGCCAAATCTAGTCCAAATCACITTCATTGATACCAT 4294
Qy 4501 TATTGTACAACTTGAGCAAGTTGCGATCAGCTCCTCAAAATGCTCTGTAAACGATG 4560
Db 4295 TATTGTACAACTTGAGCAAGTTGCGATCAGCTCCTCAAAATGCTCTGTAAACGATG 4354
Qy 4561 ACTCAACTTGACATTAACCTTTGAGCTCAGTCACTGAGTGAATGATCAGGTTGTGCA 4620
Db 4355 ACTCAACTTGACATTAACCTTTGAGCTCAGTCACTGAGTGAATGATCAGGTTGTGCA 4414
Qy 4621 GCTGTGACGACATPAGGGAACACGCGTTTCTTACCAAACTCAAGAAATTTATCAAACT 4680
Db 4415 GCTGTGACGACATPAGGGAACACGCGTTTCTTACCAAACTCAAGAAATTTATCAAACT 4474
Qy 4681 CTGCAACACTTGCATGACAGTACGAGGGAATGTCATCTTGAAGTCGACAGTGAG 4740
Db 4475 CTGCAACACTTGCATGACAGTACGAGGGAATGTCATCTTGAAGTCGACAGTGAG 4534
Qy 4741 TGTAGTCTTGAGAAATCTGAAGCGGTATTTTATATCATGAGTCACTCAGGAGGA 4800
Db 4535 TGTAGTCTTGAGAAATCTGAAGCGGTATTTTATATCATGAGTCACTCAGGAGGA 4594
Qy 4801 TCCTCTAGCGGAGCGATCTGTGGCGGA 4828
Db 4595 TCCTCTAGCGGAGCGATCTGTGGCGGA 4654
Qy 4829 4828
Db 4655 CTCGTGAAGAGGTGTTGCTGACTCATACAGGCTGAAATCGCCCCCATCATCCAGCCAGA 4714
Qy 4829 4828
Db 4715 AAGTGAGGGAGCCACGGTTGATGAGAGCTTTGTTGTAGGTGGACCAAGTTGTGATTTTGA 4774
Qy 4829 4828
Db 4775 ACTTTTGTCTTCCACGGAAACGGTCTGCTGTGTCGGGAGAGATCGGTGATCTGATCCTTCA 4834
Qy 4829 4828

Db 4835 ACTCAGCAAAAGTTCGATTTATTCAACAAAGCCGCGTCCGTCAGTCAAGTCAAGTCAATGCT 4894
Qy 4829 4828
Db 4895 CTGCCAGTGTACACCAATTAACCAATTTCTGATTAGAAAAAATCATCGAGCATCAAAATG 4954
Qy 4829 4828
Db 4955 AAATGCAATTTATTTCATATCAGGATATCAATACCATAATTTTGA AAAAGCCGTTTCTG 5014
Qy 4829 4828
Db 5015 TAATGAGGAGAAAAACTCACCGAGCAGTTCATAGGATGGCAAGATCCTGTATCGGTC 5074
Qy 4829 4828
Db 5075 TCGGATTCGACTCGTCCCAACATCAATACAACTTATTAATTTCCCTCGTCAAAAAATAAG 5134
Qy 4829 4828
Db 5135 GTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAAATGCAAAAGCTT 5194
Qy 4829 4828
Db 5195 ATGCATTTCTTTCCAGACTTGTTCACAGGCGCAGCCATTACGCTGCTCATCAAAATCACT 5254
Qy 4829 4828
Db 5255 CGCATCAACCAACCGTTATTTCATTCGTGATTGCGCTGAGGAGACGAATAACGCGATC 5314
Qy 4829 4828
Db 5315 GCTGTTAAAAAGGACAATTACAAACAGGAATCGAATGCAACCGCGCAGGAACAACGCGCAG 5374
Qy 4829 4828
Db 5375 CGCATCAACATATTTTCACTGATCAGGATATTTCTTAATACCTGGAATGCTGTTTT 5434
Qy 4829 4828
Db 5435 CCGGGGATCGCAGTGTGAGTAACCATGCATCATCAGGAGTACGATAAAATGCTTGAT 5494
Qy 4829 4828
Db 5495 GGTGGAAGAGGCATPAAATTCGTCAGCCAGTTTGTAGTCTGACCATCTCATCTGTAAACATC 5554
Qy 4829 4828
Db 5555 ATTGCAACGCTACCTTTGCCATGTTTCAGAAAAACAATCTGCGCATCGGGCTTCCCATATA 5614
Qy 4829 4828
Db 5615 CAATCGATAGATTGTGCGACCTGATTGCGGACATTAATCGGAGCCCATTTATACCCATA 5674
Qy 4829 4828
Db 5675 TAAATCAGCATCCATGTTGGAATTTAATCGCGGCTCGAGCAAGACGTTTCCCGTTGAAT 5734
Qy 4829 4828
Db 5735 ATGGCTCATAAACCCCTTGTATTACTGTTATGTATGTAAGCAGACAGTTTATTTGTTGATGA 5794
Qy 4829 4828
Db 5795 TGATATATTTTATCTTGTGCAATGTACATCAGAGATTTTGAGACACAACGTCGCTTTC 5854
Qy 4829 4828
Db 5855 CCCCCCCCCCTGACAGTCCGCATCACCGGCGCCACAGTGTGCTGGGCTTATAT 5914
Qy 4880 CCGCGACATCACCGATGGGAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGT 4939
Db 5915 CCGCGACATCACCGATGGGAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGT 5974

QY	4940	CGCGTGGGTATGTTGGCAGGCCCGCGTGGCGGGGGAATGTTGGCGCCATCTCTCTTGA	4999
Db	5975	CGCGTGGGTATGTTGGCAGGCCCGCGTGGCGGGGGAATGTTGGCGCCATCTCTCTTGA	6034
QY	5000	TGACCATTTCTTGGCGGGCGGTCTCAACGGCGCTCAACCTACTACTAGGCTGCTTCCCT	5059
Db	6035	TGACCATTTCTTGGCGGGCGGTCTCAACGGCGCTCAACCTACTACTAGGCTGCTTCCCT	6094
QY	5060	AATGAGGAGTGGCAATAAGGAGAGCGTCGAGTATCTATGATGGAAGTATGGGAATGGT	5119
Db	6095	AATGAGGAGTGGCAATAAGGAGAGCGTCGAGTATCTATGATGGAAGTATGGGAATGGT	6154
QY	5120	GATACCCGCACTTCTCAGTGTCTTGGGTCTTCTCAGTCTCTCTATCAGATATGCCCAACTAAGCAAC	5179
Db	6155	GATACCCGCACTTCTCAGTGTCTTGGGTCTTCTCAGTCTCTCTATCAGATATGCCCAACTAAGCAAC	6214
QY	5180	CGGAGGAGGAGATTTCAAGTAAATTTCTTGACCTTTTGGTCTCAGTAGACTCGAATCG	5239
Db	6215	CGGAGGAGGAGATTTCAAGTAAATTTCTTGACCTTTTGGTCTCAGTAGACTCGAATCG	6274
QY	5240	TGAGACTATCTCGGTATGACAGCAGAAATGTCCTTCTTGAGACAGTAAATGAGTCCC	5299
Db	6275	TGAGACTATCTCGGTATGACAGCAGAAATGTCCTTCTTGAGACAGTAAATGAGTCCC	6334
QY	5300	ACCAATAAGAAATCCTTGTATCAGGAACAAACTTCTTGTGAACTTTTTCGGTGCC	5359
Db	6335	ACCAATAAGAAATCCTTGTATCAGGAACAAACTTCTTGTGAACTTTTTCGGTGCC	6394
QY	5360	TGAACTATAAATAGAGTGATATGTCGGGTAGGAATCGAGCGGCCAAATCTTACC	5419
Db	6395	TGAACTATAAATAGAGTGATATGTCGGGTAGGAATCGAGCGGCCAAATCTTACC	6454
QY	5420	TTCGAGCTTCAAGAGTATGAGGTGTTGTAGATCTGATGCACTTCAAGTGAAC	5479
Db	6455	TTCGAGCTTCAAGAGTATGAGGTGTTGTAGATCTGATGCACTTCAAGTGAAC	6514
QY	5480	GTTGCTATTTCGTTCAAAACCAATTCGGAATCCAGAGAAATCAAGTGTGTTCTACTATT	5539
Db	6515	GTTGCTATTTCGTTCAAAACCAATTCGGAATCCAGAGAAATCAAGTGTGTTCTACTATT	6574
QY	5540	GATCAAGCCAGTGGGTCTTGAACCTGACAAATAGTGTGCTCGTGTGTTGAGGTCACTT	5599
Db	6575	GATCAAGCCAGTGGGTCTTGAACCTGACAAATAGTGTGCTCGTGTGTTGAGGTCACTT	6634
QY	5600	TGATGAATAAATCTAGTCTTGTGATCTAAATAATCTTGACGAGCAAGCGCATAAATACC	5659
Db	6635	TGATGAATAAATCTAGTCTTGTGATCTAAATAATCTTGACGAGCAAGCGCATAAATACC	6694
QY	5660	CAATCTAAAATCTTTTAAACCGTTAAAGGACAAAGTATGTCCTGTATTAACCC	5719
Db	6695	CAATCTAAAATCTTTTAAACCGTTAAAGGACAAAGTATGTCCTGTATTAACCC	6754
QY	5720	AAATCAGCTCGTAGTCTGATCTCATCACTTGGGGGCACTATCTTGTGTTAGAAAT	5779
Db	6755	AAATCAGCTCGTAGTCTGATCTCATCACTTGGGGGCACTATCTTGTGTTAGAAAT	6814
QY	5780	TTGCGGAGATGCAATATCAGAAAAGGTACGCTGATTTAAAGGTGAATTTATCTCAA	5839
Db	6815	TTGCGGAGATGCAATATCAGAAAAGGTACGCTGATTTAAAGGTGAATTTATCTCAA	6874
QY	5840	GATCTCTGCTCGCGCTTTCGGTGTATGACCGTGAACCTCTGACATGCACTCCG	5899
Db	6875	GATCTCTGCTCGCGCTTTCGGTGTATGACCGTGAACCTCTGACATGCACTCCG	6934
QY	5900	GAGACGCTCACAGCTTGTCTGTAGCGGATGCGGGGAGCAGCAAGCCGTCAGGGGCG	5959
Db	6935	GAGACGCTCACAGCTTGTCTGTAGCGGATGCGGGGAGCAGCAAGCCGTCAGGGGCG	6994
QY	5960	TCAGCGGTGTTGGCGGTGTCGGGCGCAGCCATGACCCAGTCACTAGCGATAGCGGA	6019
Db	6995	TCAGCGGTGTTGGCGGTGTCGGGCGCAGCCATGACCCAGTCACTAGCGATAGCGGA	7054

QY	6020	GTGTATACTGCTTAACTATGCGGCATCAGAGCAGATTTGACTGAGAGTGCAACATATGC	6079
Db	7055	GTGTATACTGCTTAACTATGCGGCATCAGAGCAGATTTGACTGAGAGTGCAACATATGC	7114
QY	6080	GGTGTGAAATACCGCAGACAGATCGTAAGGAGAAATACCGCATCAGCGCTCTTCCGCTT	6139
Db	7115	GGTGTGAAATACCGCAGACAGATCGTAAGGAGAAATACCGCATCAGCGCTCTTCCGCTT	7174
QY	6140	CCTCGCTCACTGACTCGCTCGCTCGTTCGGTTCGGTTCGGGCGAGCGGTATCAGTCACT	6199
Db	7175	CCTCGCTCACTGACTCGCTCGCTCGTTCGGTTCGGTTCGGGCGAGCGGTATCAGTCACT	7234
QY	6200	CAAGCGGTAAATACGGTTATCCACAGATCAGGGATTAACGAGGAAAGAAATGTGAG	6259
Db	7235	CAAGCGGTAAATACGGTTATCCACAGATCAGGGATTAACGAGGAAAGAAATGTGAG	7294
QY	6260	CAAAAAGCCAGCAAAAAGGCGCAGGAAACCGTAAAGAGGCGGTTGCTGGGCTTTTCCATA	6319
Db	7295	CAAAAAGCCAGCAAAAAGGCGCAGGAAACCGTAAAGAGGCGGTTGCTGGGCTTTTCCATA	7354
QY	6320	GGCTCGCCCGCTGACAGAGCATCAAAAAATCGACGCTCAAGTCAAGAGTGGCGAAACC	6379
Db	7355	GGCTCGCCCGCTGACAGAGCATCAAAAAATCGACGCTCAAGTCAAGAGTGGCGAAACC	7414
QY	6380	CGACAGACTATATAAGATACAGGCGCTTCCCGCTGGAAGCTCCCTCGTGCCTCTCTG	6439
Db	7415	CGACAGACTATATAAGATACAGGCGCTTCCCGCTGGAAGCTCCCTCGTGCCTCTCTG	7474
QY	6440	TTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTTCCTCTCGGAAAGCGTGGCGC	6499
Db	7475	TTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTTCCTCTCGGAAAGCGTGGCGC	7534
QY	6500	TTTCTCATAGCTCAGCTGTAGTATCTCAGTTTGGTGTAGTGTGCTTCCCTCCAACTGG	6559
Db	7535	TTTCTCATAGCTCAGCTGTAGTATCTCAGTTTGGTGTAGTGTGCTTCCCTCCAACTGG	7594
QY	6560	GCTGTGTGACGAAACCCCGCTTACGCGGACCGCTTATCCGCTTATCCGTAATCTCGTC	6619
Db	7595	GCTGTGTGACGAAACCCCGCTTACGCGGACCGCTTATCCGCTTATCCGTAATCTCGTC	7654
QY	6620	TTGAGTCCAAACCGGTAAAGACAGTATTCGCACTGCGAGCAGTGTGTAACAGGA	6679
Db	7655	TTGAGTCCAAACCGGTAAAGACAGTATTCGCACTGCGAGCAGTGTGTAACAGGA	7714
QY	6680	TTAGCAGAGCAGGATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCTTAACTACG	6739
Db	7715	TTAGCAGAGCAGGATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCTTAACTACG	7774
QY	6740	GCTACCTAGAAGCAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAA	6799
Db	7775	GCTACCTAGAAGCAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAA	7834
QY	6800	AAAGAGTTGGTGTCTTGTATCCGCAACAAACCAACCGCTGCTGAGCGGTGTTTTTG	6859
Db	7835	AAAGAGTTGGTGTCTTGTATCCGCAACAAACCAACCGCTGCTGAGCGGTGTTTTTG	7894
QY	6860	TTTGCAAGCAGCAGATTTACCGCAGAAAAAAGGATCTCAAGAAGATCTTTTGAATCTTT	6919
Db	7895	TTTGCAAGCAGCAGATTTACCGCAGAAAAAAGGATCTCAAGAAGATCTTTTGAATCTTT	7954
QY	6920	CTACGGGTCTGACGCTGAGTGGACGAAACTCAGGTTAAGGATTTTGGTCAAGAT	6979
Db	7955	CTACGGGTCTGACGCTGAGTGGACGAAACTCAGGTTAAGGATTTTGGTCAAGAT	8014
QY	6980	TATCAAAAAGGATCTTCACTAGATCTTTTAAATTTAAATTTAAATTTAAATCAATCT	7039
Db	8015	TATCAAAAAGGATCTTCACTAGATCTTTTAAATTTAAATTTAAATTTAAATCAATCT	8074
QY	7040	AAAGTATATAGTAAACTTGGTGTGACAGTTACCAATGCTTAATCAGTGGGACCTTA	7099
Db	8075	AAAGTATATAGTAAACTTGGTGTGACAGTTACCAATGCTTAATCAGTGGGACCTTA	8134
QY	7100	TCTCAGCATGCTCTATTTCGTTTCATCCATAGTTGCTCCCGCTGCTAGATAA	7159

Db	8135	TCTCACGCACTCTGTCTTATTTTCGTTTCATCTCATPAGTTCCTGACTCCCGCTCGTGTAGATAAA	8134
Qy	7160	CTACGATACGGGAGGGCTTACCATCTGGCCCCAGTCTGCAATATGATACCGCGAGACCCAC	7219
Db	8195	CTACGATACGGGAGGGCTTACCATCTGGCCCCAGTCTGCAATATGATACCGCGAGACCCAC	8254
Qy	7220	GCTCACGGCTCCAGATTTATCAGCAATAAACACGACGCGCGAGAGGGCCGAGCCGAGAA	7279
Db	8255	GCTCACGGCTCCAGATTTATCAGCAATAAACACGACGCGCGAGAGGGCCGAGCCGAGAA	8314
Qy	7280	GTGGTCTCTCAACTTTATCCGCTCCCATCCAGTCTAATTAATTGTTGCCGGGAAGCTAGAG	7339
Db	8315	GTGGTCTCTCAACTTTATCCGCTCCCATCCAGTCTAATTAATTGTTGCCGGGAAGCTAGAG	8374
Qy	7340	TAAGTAGTTTCGCCAGCTTAATAGTTTGGCAACGTTGTTGCCATTGCTCGAGGCATCTGG	7399
Db	8375	TAAGTAGTTTCGCCAGCTTAATAGTTTGGCAACGTTGTTGCCATTGCTCGAGGCATCTGG	8434
Qy	7400	TGTCACGCTCTGCTTGGTTGGTAGTGCCTTCAITTCAGCTCCGCTTCCCAACGATCAAGCGCAG	7459
Db	8435	TGTCACGCTCTGCTTGGTTGGTAGTGCCTTCAITTCAGCTCCGCTTCCCAACGATCAAGCGCAG	8494
Qy	7460	TTACATGATCCCCCATGTTGTGC AAAAAGGGTTAGCTCCTTCGGTCTCCGATCCGTTG	7519
Db	8495	TTACATGATCCCCCATGTTGTGC AAAAAGGGTTAGCTCCTTCGGTCTCCGATCCGTTG	8554
Qy	7520	TCAGAAGTAGTTGGCCCGCAGTGTATCACTCATGTGTTATGCGAGCACTGCATAATTCTC	7579
Db	8555	TCAGAAGTAGTTGGCCCGCAGTGTATCACTCATGTGTTATGCGAGCACTGCATAATTCTC	8614
Qy	7580	TTACTGTCATGCCATCCCTAGAGTCCTTTCTGTCGACTGGTGAAGTACTCAACCAAGTCAT	7639
Db	8615	TTACTGTCATGCCATCCCTAGAGTCCTTTCTGTCGACTGGTGAAGTACTCAACCAAGTCAT	8674
Qy	7640	TCTGAGAAATAGTGTATGCGCGCAGGAGTTGCTCTTTCGCCCGCGTCAACACGGGATAATA	7699
Db	8675	TCTGAGAAATAGTGTATGCGCGCAGGAGTTGCTCTTTCGCCCGCGTCAACACGGGATAATA	8734
Qy	7700	CCGCGCCCATAGCAGACACTTTAAAGTGCTCATCATTCGAAAACGTTCTTCGCGGGCGAA	7759
Db	8735	CCGCGCCCATAGCAGACACTTTAAAGTGCTCATCATTCGAAAACGTTCTTCGCGGGCGAA	8794
Qy	7760	AACCTCTAAGGATCTTACCGCTGTGTAGATCCAGTTTCGATGTAAACCCACTCGTGACCCCA	7819
Db	8795	AACCTCTAAGGATCTTACCGCTGTGTAGATCCAGTTTCGATGTAAACCCACTCGTGACCCCA	8854
Qy	7820	ACTGATCTTCAGCATCTTTTACTTTTCCACGAGCGTTTCTGGGTGACAAAAACAGAAAGGC	7879
Db	8855	ACTGATCTTCAGCATCTTTTACTTTTCCACGAGCGTTTCTGGGTGACAAAAACAGAAAGGC	8914
Qy	7880	AAATTCGCGCAAAAAGGGAATAAGGCGGACACGGAAATGTTGAATATCTATCTCTTCC	7939
Db	8915	AAATTCGCGCAAAAAGGGAATAAGGCGGACACGGAAATGTTGAATATCTATCTCTTCC	8974
Qy	7940	TTTTTCAATATATTGAAGCAATTTATCAGGGTTATTGCTCATGAGCGGATACATATTTG	7999
Db	8975	TTTTTCAATATATTGAAGCAATTTATCAGGGTTATTGCTCATGAGCGGATACATATTTG	9034
Qy	8000	AATGATTTAGAAAAATAACAAATAGGGTTCGGCGCACTTCCCGGAAAAGTGCAC	8059
Db	9035	AATGATTTAGAAAAATAACAAATAGGGTTCGGCGCACTTCCCGGAAAAGTGCAC	9094
Qy	8060	CTGAGCTCTAAGAAACCAATTTATCATGACATTAACCTATAAAAAATAGGCGTATCACGA	8119
Db	9095	CTGAGCTCTAAGAAACCAATTTATCATGACATTAACCTATAAAAAATAGGCGTATCACGA	9154
Qy	8120	GCGCCCTTTTCGCTTCAGAAATTAATCTCATGTTGACAGCTTATCATCTGATAGCTGAC	8179
Db	9155	GCGCCCTTTTCGCTTCAGAAATTAATCTCATGTTTGAAGCTTTATCATCTGATAGCTGAC	9214
Qy	8180	TCATGTTGGTATTGTGAATAGACGCGAGATCCGGAACTGAAAAATAACAGTTATTATT	8239

Db	9215	TCATGTGGTATTGTGAAATAGACGCAGATCGGGAACACTGAAAAATAACAGTTATTATT	9274
Qy	8240	CG 8241	
Db	9275	CG 9276	
RESULT 8			
AX207941			
LOCUS	AX207941	9276 bp	DNA
DEFINITION	Sequence 45 from Patent WO0157194.		linear
ACCESSION	AX207941		
VERSION	AX207941.1	GI:15422539	
KEYWORDS			
SOURCE	Pichia pastoris		
ORGANISM	Pichia pastoris		
REFERENCE	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Pichia.		
AUTHORS	Madison, E.L., Ong, E.O. and Yeh, J.C.		
TITLE	Nucleic acid molecules encoding transmembrane serine proteases, the encoded proteins and methods based thereon		
JOURNAL	Patent: WO 0157194-A 45 09-AUG-2001;		
FEATURES	CORVAS INTERNATIONAL, INC. (US)		
source	location/Qualifiers		
	1..9276		
	/organism="Pichia pastoris"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:4922"		
ORIGIN			
Query Match	79.6%;	Score 6561.6;	DB 6; Length 9276;
Best Local Similarity	84.7%;	Pred. NO. 0;	
Matches 8031;	Conservative 0;	Mismatches 4;	Indels 1447; Gaps 2;
Qy	1	AGATCTAACTCCAAAGACGAAAGGTTGAATGAACCTTTTGGCATCCGACATCCACAG	60
Db	1	AGATCTAACTCCAAAGACGAAAGGTTGAATGAACCTTTTGGCATCCGACATCCACAG	60
Qy	61	GTCCATTCTCACATAGTGCACCAAGCAACAGAGGGGATACACTAGCAGCAGCCGT	120
Db	61	GTCCATTCTCACATAGTGCACCAAGCAACAGAGGGGATACACTAGCAGCAGCCGT	120
Qy	121	TGCAAAACGAGGACTCCACTCTCTCTCTCTCAACACCCACTTTTGCCATCGAAAAACC	180
Db	121	TGCAAAACGAGGACTCCACTCTCTCTCTCTCAACACCCACTTTTGCCATCGAAAAACC	180
Qy	181	AGCCAGATTATGGGCTTGATTGGAGCTGCCTCATTCGATTCCTTTTATAGGCTACTA	240
Db	181	AGCCAGATTATGGGCTTGATTGGAGCTGCCTCATTCGATTCCTTTTATAGGCTACTA	240
Qy	241	ACACCATGACCTTTATAGCTGCTATCTCTGCCCCCTGCGAGGTTCACTGTTTGT	300
Db	241	ACACCATGACCTTTATAGCTGCTATCTCTGCCCCCTGCGAGGTTCACTGTTTGT	300
Qy	301	TTTCCGAATGCAACAAAGCTCCGATTTACACCCGAACTCACTCCAGATGAGGCTTTCTG	360
Db	301	TTTCCGAATGCAACAAAGCTCCGATTTACACCCGAACTCACTCCAGATGAGGCTTTCTG	360
Qy	361	AGTGGGGTCAATATAGTTTCATGTTCCCAAAATGGCCCAAACTGACAGTTAAAGCT	420
Db	361	AGTGGGGTCAATATAGTTTCATGTTCCCAAAATGGCCCAAACTGACAGTTAAAGCT	420
Qy	421	GTCTTGAACCTAATATGACAAAGCGTGATCTCATCAAGATGAACCTAGTTTGGTTCG	480
Db	421	GTCTTGAACCTAATATGACAAAGCGTGATCTCATCAAGATGAACCTAGTTTGGTTCG	480
Qy	481	TTGAAATGCTHAACGGCAGTTGCTCAAAAGAACTTCCAAAGTGCACATACGTTTGT	540
Db	481	TTGAAATGCTHAACGGCAGTTGCTCAAAAGAACTTCCAAAGTGCACATACGTTTGT	540
Qy	541	CTTGTGTTGGTATTCATTGACGAATGCTCAAAATTAATCTCATTAATGCTAGCGAGTCT	600

Db 541 CTTGTTTGGTATGATTGACGAAATGCTCAAAAATAATCTCATTAATGCTTACGGCAGTCT 600
Qy 601 CTCTATCGCTTCTGAAACCCGGTGCACTGCGGAAACGCAAAATGGGAAACACCCGCT 660
Db 601 CTCTATCGCTTCTGAAACCCGGTGCACTGCGGAAACGCAAAATGGGAAACACCCGCT 660
Qy 661 TTTTGGATGATTATGCAATGCTTCCACATTTGATGCTTCCAGATTTCTGGTGGGAATACT 720
Db 661 TTTTGGATGATTATGCAATGCTTCCACATTTGATGCTTCCAGATTTCTGGTGGGAATACT 720
Qy 721 GCTGATAGCCTAAACGCTTCATGATCAAAATTAACCTTGTTCTAACCCCTACTTGCACAGCAAT 780
Db 721 GCTGATAGCCTAAACGCTTCATGATCAAAATTAACCTTGTTCTAACCCCTACTTGCACAGCAAT 780
Qy 781 ATATAAACAGAAAGAGCTGCCGTGCTTAACCTTTTTTTTTTATATCATATATAGCTT 840
Db 781 ATATAAACAGAAAGAGCTGCCGTGCTTAACCTTTTTTTTTTATATCATATATAGCTT 840
Qy 841 ACTTTTCAATATTTGCCACTTGGTTCCAATTAACAGCTTTTGAATTTTAAAGCTTTTAAACGA 900
Db 841 ACTTTTCAATATTTGCCACTTGGTTCCAATTAACAGCTTTTGAATTTTAAAGCTTTTAAACGA 900
Qy 901 CAACTTGAGAGATCAAAAACAACTAATTTTGAAGGATCCAAAGATGAGATTTCCCT 960
Db 901 CAACTTGAGAGATCAAAAACAACTAATTTTGAAGGATCCAAAGATGAGATTTCCCT 960
Qy 961 TCAATTTTACTGCAAGTTTATTTCCGAGCATCTCCCGCATTTAGCTTCCAGTCAACACT 1020
Db 961 TCAATTTTACTGCAAGTTTATTTCCGAGCATCTCCCGCATTTAGCTTCCAGTCAACACT 1020
Qy 1021 ACAACAGAGATGAACCGCAAAATTCGGCTGAAGCTGTATCGAGTTACTCAGATTTA 1080
Db 1021 ACAACAGAGATGAACCGCAAAATTCGGCTGAAGCTGTATCGAGTTACTCAGATTTA 1080
Qy 1081 GAAGGGATTTGATGTTGCTGTTTGGCATTTTCCACAGACAAATTAACGGGTTATTG 1140
Db 1081 GAAGGGATTTGATGTTGCTGTTTGGCATTTTCCACAGACAAATTAACGGGTTATTG 1140
Qy 1141 TTTTATAATACTACTATTTGCCAGCATTTCTGCTTAAGAAAGAGGGTATCTCTCGAGAA 1200
Db 1141 TTTTATAATACTACTATTTGCCAGCATTTCTGCTTAAGAAAGAGGGTATCTCTCGAGAA 1200
Qy 1201 AGAGAGGTGAAGCCCGCCAGGCCCGAGGCCAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
Db 1201 AGAGAGGTGAAGCTTA----- 1217
Qy 1261 GCGCCCGGAGAGGATTAATTTGGCTTCCCGGTGTACCGGCCAGCAGTGCACGGAG 1320
Db 1218 ----- 1217
Qy 1321 AGAGGTGCTGTTTGTATGACAGTGTCCGGGATTTCCCGTGGTGTCTCCACCCCATGGCC 1380
Db 1218 ----- 1217
Qy 1381 ATCGAGAACACTCAAGAGAGAGATGTCCTTCTAACTAGTGGCGTAGAATTCCTTAGGG 1440
Db 1218 -----CGTAGAATTCCTTAGGG 1234
Qy 1441 CGGCGCGCAATTAATTCGCTTAGACATGACTGTTCTCAGTTCAGTTGGGCACTTACG 1500
Db 1235 CGGCGCGCAATTAATTCGCTTAGACATGACTGTTCTCAGTTCAGTTGGGCACTTACG 1294
Qy 1501 AGAAGACCGCTTCTGATGTTCTAATCAAGAGGATGTCAGATGCCATTTGCCGTGAGAG 1560
Db 1295 AGAAGACCGCTTCTGATGTTCTAATCAAGAGGATGTCAGATGCCATTTGCCGTGAGAG 1354
Qy 1561 ATGCAAGGCTTCATTTTGTATCTTTTATTTGTAACCTATATAGTATAGGATTTTTTTT 1620
Db 1355 ATGCAAGGCTTCATTTTGTATCTTTTATTTGTAACCTATATAGTATAGGATTTTTTTT 1414
Qy 1621 GTCAATTTGTTCTCTGAGAGCTTGTCTGATCAGCCTATCTCGAGCTGATGAA 1680
Db 1415 GTCAATTTGTTCTCTGAGAGCTTGTCTGATCAGCCTATCTCGAGCTGATGAA 1474

Qy 1681 TATCTTTGTTAGGGTTTGGGAAAAATCATTCGAGTTTGATGTTTCTTGTGTTATTCCT 1740
Db 1475 TATCTTTGTTAGGGTTTGGGAAAAATCATTCGAGTTTGATGTTTCTTGTGTTATTCCT 1534
Qy 1741 ACTCTCTTTCAGAGTACAGAGATTAAGTGAGAGTTTCTGTCGCAAGCTTATCCGATAA 1800
Db 1535 ACTCTCTTTCAGAGTACAGAGATTAAGTGAGAGTTTCTGTCGCAAGCTTATCCGATAA 1594
Qy 1801 GCTTTAATGCGTATGTTTATACAGTTAAATTTGCTAAACAGTCAAGCACCGGTATGAA 1860
Db 1595 GCTTTAATGCGTATGTTTATACAGTTAAATTTGCTAAACAGTCAAGCACCGGTATGAA 1654
Qy 1861 ATCTAAATGCGCTCATCTGTCATCTCGCACCGCTCACCTGGATGCTGTAGGATAGG 1920
Db 1655 ATCTAAATGCGCTCATCTGTCATCTCGCACCGCTCACCTGGATGCTGTAGGATAGG 1714
Qy 1921 CTTGGTTATGCGCGGTACTGCGCGGCTCTTGGGGATATCGTCCATTCGACAGCATCGC 1980
Db 1715 CTTGGTTATGCGCGGTACTGCGCGGCTCTTGGGGATATCGTCCATTCGACAGCATCGC 1774
Qy 1981 CAGTCACTATGCGGTGCTGCTAGCTATATGCGTTGATGCAATTTCTATGCGCACCGGT 2040
Db 1775 CAGTCACTATGCGGTGCTGCTAGCTATATGCGTTGATGCAATTTCTATGCGCACCGGT 1834
Qy 2041 TCTCGAGACACTGTCGACCGCTTTGGCGCGCCCGCAGTCTGCTCGCTTCGCTACTTGG 2100
Db 1835 TCTCGAGACACTGTCGACCGCTTTGGCGCGCCCGCAGTCTGCTCGCTTCGCTACTTGG 1894
Qy 2101 AGCCACTATGACTACCGGATCATGCGACCAACACCCGCTCTGTGGATCTATCGAATCTA 2160
Db 1895 AGCCACTATGACTACCGGATCATGCGACCAACACCCGCTCTGTGGATCTATCGAATCTA 1954
Qy 2161 AATGTAAGTTAAATCTCTAAATAAATAAATTAAGTCCCAAGTTCTCCATACGAACTTAA 2220
Db 1955 AATGTAAGTTAAATCTCTAAATAAATAAATTAAGTCCCAAGTTCTCCATACGAACTTAA 2014
Qy 2221 CAGCATTTGCGGTGAGCATCTAGACCTTTCAACAGCAGCAGATCCATCCTGCTTGGCCAA 2280
Db 2015 CAGCATTTGCGGTGAGCATCTAGACCTTTCAACAGCAGCAGATCCATCCTGCTTGGCCAA 2074
Qy 2281 TATGTTTCACTTCCCTCAGAGTTTACGCTTGTGAGTGTAACTTCTGGAAGTTGCGAG 2340
Db 2075 TATGTTTCACTTCCCTCAGAGTTTACGCTTGTGAGTGTAACTTCTGGAAGTTGCGAG 2134
Qy 2341 TGTAACTCCGCTGATTTGACGGGCATATCCGTCAGTTCGGCAAGTGTGTTGGTACCGG 2400
Db 2135 TGTAACTCCGCTGATTTGACGGGCATATCCGTCAGTTCGGCAAGTGTGTTGGTACCGG 2194
Qy 2401 AGAGTAAATCTCCAACTCTCTGGAGATGAGCACCAACAAACAGATCCAGCGTGT 2460
Db 2195 AGAGTAAATCTCCAACTCTCTGGAGATGAGCACCAACAAACAGATCCAGCGTGT 2254
Qy 2461 GTACTTTGATCAACATGAAGAGAGCATTTCTCGATTTGCAAGGATCAAGTCTTCAGAGCGT 2520
Db 2255 GTACTTTGATCAACATGAAGAGAGCATTTCTCGATTTGCAAGGATCAAGTCTTCAGAGCGT 2314
Qy 2521 ACTGATTTGACATTTTCCAAAGCCTGCTGCTAGTTTGCACCCGATAGGGTTGTAGAGTGT 2580
Db 2315 ACTGATTTGACATTTTCCAAAGCCTGCTGCTAGTTTGCACCCGATAGGGTTGTAGAGTGT 2374
Qy 2581 CAATACACTTGGTACAAATTTCAACCTTTGGCAACATGACAGCTTGGTGTGAAAGCAT 2640
Db 2375 CAATACACTTGGTACAAATTTCAACCTTTGGCAACATGACAGCTTGGTGTGAAAGCAT 2434
Qy 2641 CTTCAATTTGCAAGCTCTCTGCTGCTCATATCGACAGCCCAACAGAAATCACTGGGAAT 2700
Db 2435 CTTCAATTTGCAAGCTCTCTGCTGCTCATATCGACAGCCCAACAGAAATCACTGGGAAT 2494
Qy 2701 CAATACACTTGGTGTGAGACAGAGGTCTGAGGCAACGAAATCTGGATCAGCGTAT 2760
Db 2495 CAATACACTTGGTGTGAGACAGAGGTCTGAGGCAACGAAATCTGGATCAGCGTAT 2554

2761 TATCAGCAATACTAGAACTTCAGAGGGCCAGAGGCATGTCAATACATACAGGGCTG 2820
 Db
 2555 TATCAGCAATACTAGAACTTCAGAGGGCCAGAGGCATGTCAATACATACAGGGCTG 2614
 Qy
 2821 ATGTGTCATTTTGAACCATCATCTTGGCAGCAATACGAACTGGTTCTTGGACCAATA 2880
 Db
 2615 ATGTGTCATTTTGAACCATCATCTTGGCAGCAATACGAACTGGTTCTTGGACCAATA 2674
 Qy
 2881 TTTTGTGACACTTAGGAACAGTTTCTGTTCCGTAAGCCATAGCAGCTACTGCTCGGGCG 2940
 Db
 2675 TTTTGTGACACTTAGGAACAGTTTCTGTTCCGTAAGCCATAGCAGCTACTGCTCGGGCG 2734
 Qy
 2941 CTCCTGCTAGCAGCATACCTTAGCACCACCACTTGTGGGCAACGTAGACTTCTGGGG 3000
 Db
 2735 CTCCTGCTAGCAGCATACCTTAGCACCACCACTTGTGGGCAACGTAGACTTCTGGGG 2794
 Qy
 3001 TAAGGTTACCATCTTCTTAGTGGAGATGCAAAACAAATTTCTTGCACACAGCAACT 3060
 Db
 2795 TAAGGTTACCATCTTCTTAGTGGAGATGCAAAACAAATTTCTTGCACACAGCAACT 2854
 Qy
 3061 TGGCAGGAACACCCAGCATCAGGGAAGTGAAGGAGAAATGGCGTTTCCACCAAGGAAT 3120
 Db
 2855 TGGCAGGACACCCAGCATCAGGGAAGTGAAGGAGAAATGGCGTTTCCACCAAGGAAT 2914
 Qy
 3121 AGAGGCCAATTTCTCAATAGTCTTGCAGGAGAGAGAGAGAGAGAGAGAGAGAGTCT 3180
 Db
 2915 AGAGGCCAATTTCTCAATAGTCTTGCAGGAGAGAGAGAGAGAGAGAGAGAGAGTCT 2974
 Qy
 3181 CAACCTTGAAGCTCTCGTTAGTTGAGCTTCAATGGAATTTCTGACGTTATCTATAGAGA 3240
 Db
 2975 CAACCTTGAAGCTCTCGTTAGTTGAGCTTCAATGGAATTTCTGACGTTATCTATAGAGA 3034
 Qy
 3241 GATCAATGGCTCTCTTAACGTTATCTGCAATTCGATAAGTTCCTCTGGGAAGGAGCTT 3300
 Db
 3035 GATCAATGGCTCTCTTAACGTTATCTGCAATTCGATAAGTTCCTCTGGGAAGGAGCTT 3094
 Qy
 3301 CTACACAGGTTCTTCAAGGAGTCCATCAACACTTGGAGTTAGTTCCTTAAAGGGCTT 3360
 Db
 3095 CTACACAGGTTCTTCAAGGAGTCCATCAACACTTGGAGTTAGTTCCTTAAAGGGCTT 3154
 Qy
 3361 TGTCAACATTTTGAAGCAATTCGCAATTCGCAATTCGCAATTCGCAATTCGCAATTCG 3420
 Db
 3155 TGTCAACATTTTGAAGCAATTCGCAATTCGCAATTCGCAATTCGCAATTCGCAATTCG 3214
 Qy
 3421 TTTTCTGATAGGAGAGAGAGGATCTTCAATTTCTTGGAGGAGGCTTGAAGAGCT 3480
 Db
 3215 TTTTCTGATAGGAGAGAGGATCTTCAATTTCTTGGAGGAGGCTTGAAGAGCT 3274
 Qy
 3481 CAATTTTGCATTAATCAATCAAGCTTCAGAGGAGCTTCTTTAGGTTTGGATTTCTTCT 3540
 Db
 3275 CAATTTTGCATTAATCAATCAAGCTTCAGAGGAGCTTCTTTAGGTTTGGATTTCTTCT 3334
 Qy
 3541 TAGGTTTCTTGGTTGATCTTGGCTTGGATCTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3600
 Db
 3335 TAGGTTTCTTGGTTGATCTTGGCTTGGATCTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3394
 Qy
 3601 ACTTCATATCCAGGTTTCTTCCACTCTGCTCAAGCTCAGACCGTACATCTGGCAGATCTAA 3660
 Db
 3395 ACTTCATATCCAGGTTTCTTCCACTCTGCTCAAGCTCAGACCGTACATCTGGCAGATCTAA 3454
 Qy
 3661 CTAATGCAAAATAAATAAGTACAGCATTCCTAGGCTATATCTTCTTGGATTTAGCTT 3720
 Db
 3455 CTAATGCAAAATAAATAAGTACAGCATTCCTAGGCTATATCTTCTTGGATTTAGCTT 3514
 Qy
 3721 CTGCAAGTTTCACTAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3780
 Db
 3515 CTGCAAGTTTCACTAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3574
 Qy
 3781 ACCGTTTGGTATAGAACCTTCTGAGCAATTCCTTACGATCCACAGGTTGGCTTCCA 3840
 Db
 3575 ACCGTTTGGTATAGAACCTTCTGAGCAATTCCTTACGATCCACAGGTTGGCTTCCA 3634
 Qy
 3841 TGGCTCTAAGACCCCTTTGATTTGGCCAAACAGGAAGTGGCTTCCAGGTGACAGAAACCA 3900

3635 TGGCTCTAAGACCCCTTTGATTTGGCCAAACAGGAAGTGCCTTCCAGTGCAGAAACCA 3694
 Qy
 3901 CACTGTTTGTTCACACCAAAATTTCAAGCAGTCTCCATCACAATCCAAATTCGATACCCA 3960
 Db
 3695 CACTGTTTGTTCACACCAAAATTTCAAGCAGTCTCCATCACAATCCAAATTCGATACCCA 3754
 Qy
 3961 GCAACTTTTGTAGTTGCTTCCAGATAGCACTTTATACCAAAACCCCTGACGAGGAT 4020
 Db
 3755 GCAACTTTTGTAGTTGCTTCCAGATAGCACTTTATACCAAAACCCCTGACGAGGAT 3814
 Qy
 4021 GGTAGCTCCAGTTTGTGCTTATAGCTCCGGAATAGACTTTTGGAGAGTACACCA 4080
 Db
 3815 GGTAGCTCCAGTTTGTGCTTATAGCTCCGGAATAGACTTTTGGAGAGTACACCA 3874
 Qy
 4081 GGGCCAAACGAGTAATTAAGAGTACGCCAACCAAGTAGTGAATAGACCATCGGGGGCT 4140
 Db
 3875 GGGCCAAACGAGTAATTAAGAGTACGCCAACCAAGTAGTGAATAGACCATCGGGGGCT 3934
 Qy
 4141 CAGTAGTCAGAGCGCCAAACAAATTTTCACTGACAGGAACTTTTTCATCTTTCAGAAA 4200
 Db
 3935 CAGTAGTCAGAGCGCCAAACAAATTTTCACTGACAGGAACTTTTTCATCTTTCAGAAA 3994
 Qy
 4201 GTTCGTATTTCACTAGTCAATTTGCCAGCATCAATTAATGGGATTTATACCAAGAACAG 4260
 Db
 3995 GTTCGTATTTCACTAGTCAATTTGCCAGCATCAATTAATGGGATTTATACCAAGAACAG 4054
 Qy
 4261 TGGAGTCAATCTACCAACTTTTGGCTCTCAGAAAAGCATTAACAGTTCTTACTACGC 4320
 Db
 4055 TGGAGTCAATCTACCAACTTTTGGCTCTCAGAAAAGCATTAACAGTTCTTACTACGC 4114
 Qy
 4321 CATTAGTCAAACTTTTCAATTCGCCAGTGGAGAGAAAAGGACGAGATCTAGCAT 4380
 Db
 4115 CATTAGTCAAACTTTTCAATTCGCCAGTGGAGAGAAAAGGACGAGATCTAGCAT 4174
 Qy
 4381 TAGCGGCAAGATGCAACTTTTATCAACAGGTCCTATAGATAACCTTAGCGCTGGGA 4440
 Db
 4175 TAGCGGCAAGATGCAACTTTTATCAACAGGTCCTATAGATAACCTTAGCGCTGGGA 4234
 Qy
 4441 TCATCTTTTGGACAACTCTTTCTGCAAACTTAGTTCAAAATCACTTCAATGATACCAT 4500
 Db
 4235 TCATCTTTTGGACAACTCTTTCTGCAAACTTAGTTCAAAATCACTTCAATGATACCAT 4294
 Qy
 4501 TATTGTACAACTTTGACAAAGTTTGTGATCAGTCTCTCAAAATTTGGTCTCTGTAAGGATG 4560
 Db
 4295 TATTGTACAACTTTGACAAAGTTTGTGATCAGTCTCTCAAAATTTGGTCTCTGTAAGGATG 4354
 Qy
 4561 ACTCAACTTGCACATTAATTTGAAGCTCAGTCTGATTAAGTGAATTTGATCAGGTTGTGA 4620
 Db
 4355 ACTCAACTTGCACATTAATTTGAAGCTCAGTCTGATTAAGTGAATTTGATCAGGTTGTGA 4414
 Qy
 4621 GCTGTGACAGCATAGGAAAACACGGCTTTTCTTACCAAACTCAAGAAATTTATCAAACT 4680
 Db
 4415 GCTGTGACAGCATAGGAAAACACGGCTTTTCTTACCAAACTCAAGAAATTTATCAAACT 4474
 Qy
 4681 CTGCAACACTTGGTATGAGGATAGGAGGAAATGTATCTTCAAGTTCGAGAGTGAAG 4740
 Db
 4475 CTGCAACACTTGGTATGAGGATAGGAGGAAATGTATCTTCAAGTTCGAGAGTGAAG 4534
 Qy
 4741 TGTAGTCTTGAAGAAATTTCTGAAGCGGTATTTTATTTATCAGTGAAGTCAATCAGGAGA 4800
 Db
 4535 TGTAGTCTTGAAGAAATTTCTGAAGCGGTATTTTATTTATCAGTGAAGTCAATCAGGAGA 4594
 Qy
 4801 TCTCTACGCCGAGCATCTGTCGCGA----- 4828
 Db
 4595 TCTCTACGCCGAGCATCTGTCGCGA----- 4654
 Qy
 4829 ----- 4828
 Db
 4655 CTGCTGAAGAGGTTTGTGCTGACTCATACAGGCTGAATGCCCCATCATCCAGCCAGA 4714
 Qy
 4829 ----- 4828

Db	4715	AAGTGAAGGACCCACGGTTGATGAGAGCTTTGTTGTTAGTGGACCACTTGGTGATTTTGA	4774
Qy	4829	-----	4828
Db	4775	ACTTTTGGCTTGGCACGGAACGGTCTGGGTTGTCGGGAAGATGGGTGATCTGATCCTTCA	4834
Qy	4829	-----	4828
Db	4835	ACTCAGCAAAAGTTTCGATTTATTCAACAAAGCCCGCTCCCGTCAAGTCAGCGTAATGCT	4894
Qy	4829	-----	4828
Db	4895	CTGCCAGTGTTACAAACCAATTAAACCAATCTGATTTAGAAAACCTCATCGAGCATCAATG	4954
Qy	4829	-----	4828
Db	4955	AAACTGCAATTTATTTCATATCAGGATTTATCAATACCATATTTTGAAGAAAGCCGTTTCTG	5014
Qy	4829	-----	4828
Db	5015	TAATGAAGGAGAAAACCTCACCAGGACGATTCATAGGATGCCAAGATCCTGGTATCGGTC	5074
Qy	4829	-----	4828
Db	5075	TGCGATTCCGACTCGTCCAAACATCAATACACCTATTAAATTTCCCTCGTCAAAAATAAG	5134
Qy	4829	-----	4828
Db	5135	GTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAAATGCCAAAAGCTT	5194
Qy	4829	-----	4828
Db	5195	ATGCATTTCTTCCAGACTTGTTCACAGGCCAGCCATTACGCTCGTCATCAAAATCACT	5254
Qy	4829	-----	4828
Db	5255	CGCATCAACCAACCGTTATTTCATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGATC	5314
Qy	4829	-----	4828
Db	5315	GCTGTTAAAGGACAATTACAAACAGGAATCGAATGCAACCGCGCGAGAACACTGGCAG	5374
Qy	4829	-----	4828
Db	5375	CCCATCAACAAATTTTCACTCGAATCAGGATATTTCTTAATCTCGAATGCTGTTTT	5434
Qy	4829	-----	4828
Db	5435	CCCGGGATCGCAGTGTGAGTAACCATGCATCATCAGGAGTACGGATPAAATGCTTGAT	5494
Qy	4829	-----	4828
Db	5495	GGTCGAGAGGCAATAATTCGTCAGCCAGTTTAGTCTGACCATTCTCATCTGTAAACATC	5554
Qy	4829	-----	4828
Db	5555	ATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCAT	5614
Qy	4829	-----	4828
Db	5615	CAATCATAGATGTTCGACCTGATGTCGCCGACATTAATCGGAGCCCATTTATACCCATA	5674
Qy	4829	-----	4828
Db	5675	TAAATCAGCATCCATGTTGGAATTTTAATCGCGCCTCGAGCAAGACGTTTCCCGTTGAAT	5734
Qy	4829	-----	4828
Db	5735	ATGGCTCATACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTATTTGTTCAATG	5794
Qy	4829	-----	4828
Db	5795	TGATATATTTTATCTTGTGCAATGTAACATCAGAGATTTTTCAGACACACGTCGCTTTC	5854

Qy	4829	-----	4828
Db	5855	CCCCCCCCCTGCAAGTCCGATCACCGCGCCACAGTGGTGGTGGCGCTATAT	5914
Qy	4880	CGCCGACATCACCGATGGGAAGATCGGCTCGCCACTTCGCGCTCATCAGCGCTGTTTT	4939
Db	5915	CGCCGACATCACCGATGGGAAGATCGGCTCGCCACTTCGCGCTCATCAGCGCTGTTTT	5974
Qy	4940	CGCGTGGGTATGTTGGCAGGCGCCGTCGCCCGGGGACTGTTGGCGGCATCTCCTTGA	4999
Db	5975	CGCGTGGGTATGTTGGCAGGCGCCGTCGCCCGGGGACTGTTGGCGGCATCTCCTTGA	6034
Qy	5000	TGCACCATTTCTTGGCGCGGTCGTCGACCGGCTCAACCTACTACTGGGCTGCTTCT	5059
Db	6035	TGCACCATTTCTTGGCGCGGTCGTCGACCGGCTCAACCTACTACTGGGCTGCTTCT	6094
Qy	5060	AATGCAGGAGTGCATTAAGGAGAGCGTCGAGTATCTATGATTGGAAGTATGGGAATGCT	5119
Db	6095	AATGCAGGAGTGCATTAAGGAGAGCGTCGAGTATCTATGATTGGAAGTATGGGAATGCT	6154
Qy	5120	GATACCCGATTTCTTCAGTGTCTGAGTCTCTCTATCAGATTTATGCCCACTAAAGCAAC	5179
Db	6155	GATACCCGATTTCTTCAGTGTCTCTCTATCAGATTTATGCCCACTAAAGCAAC	6214
Qy	5180	CGGAGGAGGAGATTTTCATGTAATTTCTCTGACTTTTGGTCTCATCAGTAGACTCGAACTG	5239
Db	6215	CGGAGGAGGAGATTTTCATGTAATTTCTCTGACTTTTGGTCTCATCAGTAGACTCGAACTG	6274
Qy	5240	TGAGACTATCTCGGTTATCAGCAGAAATGTCCTTCTTGGAGACAGTAAATGAACTCC	5299
Db	6275	TGAGACTATCTCGGTTATCAGCAGAAATGTCCTTCTTGGAGACAGTAAATGAACTCC	6334
Qy	5300	ACCAATAAGAAATCTTGTATCAGGAACAACTTCTTGTTCGAACTTTTTCGGTGGC	5359
Db	6335	ACCAATAAGAAATCTTGTATCAGGAACAACTTCTTGTTCGAACTTTTTCGGTGGC	6394
Qy	5360	TTGAACTATAAAATGTAGAGTGGATATGCGGTAGGAAATGGAGCGGCAAAATGCTTACC	5419
Db	6395	TTGAACTATAAAATGTAGAGTGGATATGCGGTAGGAAATGGAGCGGCAAAATGCTTACC	6454
Qy	5420	TTCTGACCTTCAAGAGTATGTAGGTTTGTAGTACTGATGCCAACTTCAGTGACAC	5479
Db	6455	TTCTGACCTTCAAGAGTATGTAGGTTTGTAGTACTGATGCCAACTTCAGTGACAC	6514
Qy	5480	GTTGCTATTTCGTTCAACCATTCGAAATCCAGAGAAATCAAAGTTGTTTGTCTACTAT	5539
Db	6515	GTTGCTATTTCGTTCAACCATTCGAAATCCAGAGAAATCAAAGTTGTTTGTCTACTAT	6574
Qy	5540	GATCCAGCAGTGGGCTGTTGAACTGACATAGTGTGCTGTGTTTGGAGTCACTTT	5599
Db	6575	GATCCAGCAGTGGGCTGTTGAACTGACATAGTGTGCTGTGTTTGGAGTCACTTT	6634
Qy	5600	TGATGAATAAATCTAGTCTTTGATCTAAATATCTTGACGCCCAAGCGATAAATACC	5659
Db	6635	TGATGAATAAATCTAGTCTTTGATCTAAATATCTTGACGCCCAAGCGATAAATACC	6694
Qy	5660	CAATCTAAACCTTTTAAACGTTTAAAGGACAAAGTATCTGCTGCTGTTTAAACCC	5719
Db	6695	CAATCTAAACCTTTTAAACGTTTAAAGGACAAAGTATCTGCTGCTGTTTAAACCC	6754
Qy	5720	AAATCAGCTGTTAGTCTCTCATCACTTCAGCGGCACTATCTGTTTGTAGAGAAAT	5779
Db	6755	AAATCAGCTGTTAGTCTCTCATCACTTCAGCGGCACTATCTGTTTGTAGAGAAAT	6814
Qy	5780	TTGCGAGATGCGATATCGAGAAAAAGGTACGCTGATTTTAAACGTTAAATTTATCTCAA	5839
Db	6815	TTGCGAGATGCGATATCGAGAAAAAGGTACGCTGATTTTAAACGTTAAATTTATCTCAA	6874
Qy	5840	GATCTCTGCTCGCGCTTTCGGTGATGACGGTGAACCTCTGACACATGCGGCTCCCG	5899
Db	6875	GATCTCTGCTCGCGCTTTCGGTGATGACGGTGAACCTCTGACACATGCGGCTCCCG	6934

Db 9095 CTGAGCTCTAAGAAACCAATTATTATCATGACATTAACTATATAAAATAGCGGTATCACGA 9154
Qy 8120 GGCCTCTCGCTTCAAGAAATTAATCTCATGTTTGACAGCTTATCATCGATAAGCTGAC 8179
Db 9155 GGCCTCTCGCTTCAAGAAATTAATCTCATGTTTGACAGCTTATCATCGATAAGCTGAC 9214
Qy 8180 TCATGTTGGTATTGTGAATAGACGAGATCGGGAACACATGAAATAACAGTATTATT 8239
Db 9215 TCATGTTGGTATTGTGAATAGACGAGATCGGGAACACATGAAATAACAGTATTATT 9274
Qy 8240 CG 8241
Db 9275 CG 9276
RESULT 9
LOCUS AX146618 9776 bp DNA linear PAT 31-MAY-2001
DEFINITION Sequence 80 from Patent WO0134654.
ACCESSION AX146618
VERSION AX146618.1 GI:14285011
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Strauch, K.
TITLE Hedgehog fusion proteins and uses
JOURNAL Patent: WO 0134654-A 80 17-MAY-2001;
BIOGEN, INC. (US)
FEATURES
source 1..976
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 77.5%; Score 6383.8; DB 6; Length 9776;
Best Local Similarity 82.9%; Pred. No. 0;
Matches 8101; Conservative 0; Mismatches 137; Indels 1535; Gaps 10;
Qy 2 GATCTAACATCCAAAGAGCAAGGTTGAATGAACCTTTTGCCATCCGACATCCACAGG 61
Db 1 GATCTAACATCCAAAGAGCAAGGTTGAATGAACCTTTTGCCATCCGACATCCACAGG 60
Qy 62 TCATCTCACAATAAGTGCACAAAGCAACAGGAGGGGATACACTAGCAGCAGACCGTT 121
Db 61 TCATCTCACAATAAGTGCACAAAGCAACAGGAGGGGATACACTAGCAGCAGACCGTT 120
Qy 122 GCAACGCGAGGACCTCCACTCTCTCTCTCAACACCCACCTTTTGCCATCGAAAAACA 181
Db 121 GCAACGCGAGGACCTCCACTCTCTCTCTCAACACCCACCTTTTGCCATCGAAAAACA 180
Qy 182 GCCAGTTATTGGCTTGATTGGAGTGGCTCATTTCCAAATTCCTTCTATTAGGCTACTAA 241
Db 181 GCCAGTTATTGGCTTGATTGGAGTGGCTCATTTCCAAATTCCTTCTATTAGGCTACTAA 240
Qy 242 CACCATGCTTTATTAGCTGCTCTATCCGCGCCCTCGGCGAGGTTTCAATGTTGTTTAT 301
Db 241 CACCATGCTTTATTAGCTGCTCTATCCGCGCCCTCGGCGAGGTTTCAATGTTGTTTAT 300
Qy 302 TTCGGAATGCAACAAGCTCCGATTAACCCGAAACATCACTCCAGATGAGGCTTTCTGA 361
Db 301 TTCGGAATGCAACAAGCTCCGATTAACCCGAAACATCACTCCAGATGAGGCTTTCTGA 360
Qy 362 GTGTGGGTCAATAGTTTCATGTTCCCAATGGCCCAAACTGACAGTTTAAACGCTG 421
Db 361 GTGTGGGTCAATAGTTTCATGTTCCCAATGGCCCAAACTGACAGTTTAAACGCTG 420
Qy 422 TCTTGAACTTAATGACAAAGCGTGTCTCATCCAAAGATGAACCTAAAGTTGGTTGCT 481
Db 421 TCTTGAACTTAATGACAAAGCGTGTCTCATCCAAAGATGAACCTAAAGTTGGTTGCT 480

Qy 482 TGAATGCTAAGCGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCCCATACCGTTTCTC 541
Db 481 TGAATGCTAAGCGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCCCATACCGTTTCTC 540
Qy 542 TTGTTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAAATGCTTAGCGCAGTCTC 601
Db 541 TTGTTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAAATGCTTAGCGCAGTCTC 600
Qy 602 TCTATCGCTTCTGAACCCCGGTGCACCTGTGCGGAAACGCAATGGGAAACACCCGCTT 661
Db 601 TCTATCGCTTCTGAACCCCGGTGCACCTGTGCGGAAACGCAATGGGAAACACCCGCTT 660
Qy 662 TTTGGATGATTATGCAATGCTCTCCACATGTATGCTTCCAAAGATTCTGTGGGAATACTG 721
Db 661 TTTGGATGATTATGCAATGCTCTCCACATGTATGCTTCCAAAGATTCTGTGGGAATACTG 720
Qy 722 CTGATAGCTTAACGTTTATGATCAAAATTAATTAATGTTCTTAACCCCTACTTGACAGAATA 781
Db 721 CTGATAGCTTAACGTTTATGATCAAAATTAATTAATGTTCTTAACCCCTACTTGACAGAATA 780
Qy 782 TATAACAGAGGAGTGCCTGCTTAAACCTTTTATCATCAATTAATAGCTTA 841
Db 781 TATAACAGAGGAGTGCCTGCTTAAACCTTTTATCATCAATTAATAGCTTA 840
Qy 842 CTTTCAATAATTCGACTGGTTCCAAATTGACAAGCTTTTGAATTTAAACGCTTTTAAACGAC 901
Db 841 CTTTCAATAATTCGACTGGTTCCAAATTGACAAGCTTTTGAATTTAAACGCTTTTAAACGAC 900
Qy 902 AACTGAGAGATCAAAAACAACTTAATTTTCGAGAGTCCAAACGATGAGATTTCCTT 961
Db 901 AACTGAGAGATCAAAAACAACTTAATTTTCGAGAGTCCAAACGATGAGATTTCCTT 960
Qy 962 CAATTTTACTGCACTTTTATTCGAGCATCTCTCCGATTTAGCTGCTCCAGTCAACACTA 1021
Db 961 CAATTTTACTGCACTTTTATTCGAGCATCTCTCCGATTTAGCTGCTCCAGTCAACACTA 1020
Qy 1022 CAACAGAGATGAACCGCACAAATTCGCGCTGAAGCTGTCATCGCTTACTCAGATTAG 1081
Db 1021 CAACAGAGATGAACCGCACAAATTCGCGCTGAAGCTGTCATCGCTTACTCAGATTAG 1080
Qy 1082 AAGGGATTTCGATGTTGCTGTTTTCGCAATTTTCCAAACAGCACAATAACGGGTATTGT 1141
Db 1081 AAGGGATTTCGATGTTGCTGTTTTCGCAATTTTCCAAACAGCACAATAACGGGTATTGT 1140
Qy 1142 TTATAAATACTACTATTGCCAGCATTTGCTGCTAAAGAAAGAGGGTATCTCTCGAGAAA 1201
Db 1141 TTATAAATACTACTATTGCCAGCATTTGCTGCTAAAGAAAGAGGGTATCTCTCGAGAAA 1200
Qy 1202 GAGAGCTGAAGCCCGAGG----- 1219
Db 1201 GATCGGACCGGCGAGGGGTTCCGGAAGAGGAGGCCACCCCAAAAGCTGACCCCTTTAG 1260
Qy 1220 -----CCAGGCCCGAGGCCAGGCCAG 1242
Db 1261 CTTACAGCAGTTTATCCCAATGTGGCCGAGAAGACCTTAGCGCCAGCGGAAGGTATG 1320
Qy 1243 GAAGAAACAT----- 1252
Db 1321 AAGGGAAGATCTCCAGAACTCCGAGCGATTTAAGGAATCAACCCCAATTAACAACCCG 1380
Qy 1253 -----GTATCATGGCCCCCGGAGAGGATAAATTTGTGGCTTCCCGGTGTC 1299
Db 1381 ACATCATATTTAAGGATGAAGAAAACACCGAGCGGACAGGCTGATGACTCAGAGGTGTA 1440
Qy 1300 AC-----CGCCAGCAGTGCAGGAGAGAGGTTGCTGTTT 1334
Db 1441 AGGACAAGTTGAACGCTTTTGGCCATCTCGGTGATGAACAGTGGCCAGGAGTGAACCTGC 1500
Qy 1335 TGATGACAGTGTCCGGGATTCGCGTG----- 1362
Db 1501 GGGTGACCGAGGCTGGGACGAAGATGGCCACCACTCAGAGGAGTCTCTGCACCTACAGG 1560

QY 1363 ----- 1362
Db 1561 GCCGGCAGTGCACATACACAGCTCTGACCCGAGCCGAGCAAGTACGGCATGCTGGCCCC 1620
QY 1363 -----TGCTTCCACCCCATGGCCATCGAGAACACACTCAAGAGAAGAAT 1405
Db 1621 GCCTGGCGGTGGAGCCCGCTTCGACTGGGTGTACTACGAGTCCAGGCAATATCCACT 1680
QY 1406 GTCCCTTCTACTA-----GTGGGTAGATTCCTTAGG-----GCCGCCGGA 1449
Db 1681 GCTCGGTGAAGACGAGAACTCGGTGGCGGCCAAATCGGAGGCTGATTCGGCGCCGGA 1740
QY 1450 ATTAATTCGCCCTTAGACATGACTCTTCCTCAGTTCAAGTTGGGCACATTACGAGAAGACCG 1509
Db 1741 ATTAATTCGCCCTTAGACATGACTCTTCCTCAGTTCAAGTTGGGCACATTACGAGAAGACCG 1800
QY 1510 GTCTTGCTAGATTTCTAATCAAGAGATGTGAGATGCAATTTGCTCGAGAGATGCAGGCT 1569
Db 1801 GTCTTGCTAGATTTCTAATCAAGAGATGTGAGATGCAATTTGCTCGAGAGATGCAGGCT 1860
QY 1570 TCATTTTGTACTTTTTTATTTGTAACCTATATAGTATAGGATTTTTTTTGTGCAATTTG 1629
Db 1861 TCATTTTGTACTTTTTTATTTGTAACCTATATAGTATAGGATTTTTTTTGTGCAATTTG 1920
QY 1630 TTTCTTCTCGTACGAGCTTGCTCTGATCAGCTATCTCGAGCTGATGAATATCTTGTG 1689
Db 1921 TTTCTTCTCGTACGAGCTTGCTCTGATCAGCTATCTCGAGCTGATGAATATCTTGTG 1980
QY 1690 GTAGGGTTTCGGAATATCATTGAGTTGATGTTTTTCTTGTGATTTCCCATCTCCTCTT 1749
Db 1981 GTAGGGTTTCGGAATATCATTGAGTTGATGTTTTTCTTGTGATTTCCCATCTCCTCTT 2040
QY 1750 CAGAGTACAGAAGATTAAGTCAGAAGTTGTTTGTGCAAGCTTATCGATAGCTTTAATG 1809
Db 2041 CAGAGTACAGAAGATTAAGTCAGAAGTTGTTTGTGCAAGCTTATCGATAGCTTTAATG 2100
QY 1810 CGGTAGTTTATCAGAGTTAAATCTTACGAGTCAGGCAAGCTGATGAATCTACAA 1869
Db 2101 CGGTAGTTTATCAGAGTTAAATCTTACGAGTCAGGCAAGCTGATGAATCTACAA 2160
QY 1870 TCGCTCATCTGCTATCTCGGACCCGCTCACCCTGGATCTGTAGGCATAGGCTTGGTTAT 1929
Db 2161 TGGCTCATCTGCTATCTCGGACCCGCTCACCCTGGATCTGTAGGCATAGGCTTGGTTAT 2220
QY 1930 GCCGTACTGCGGGCTCTTTGCGGGATATGTCCTCATTCGACAGCATCGCCAGTACTA 1989
Db 2221 GCCGTACTGCGGGCTCTTTGCGGGATATGTCCTCATTCGACAGCATCGCCAGTACTA 2280
QY 1990 TGGCTGCTGCTAGCGCTATATGCTGTGATGCAATTTCTATGCGCACCCGCTTCTCGGAGC 2049
Db 2281 TGGCTGCTGCTAGCGCTATATGCTGTGATGCAATTTCTATGCGCACCCGCTTCTCGGAGC 2340
QY 2050 ACTGTCGACCGCTTTGCGCCGCCGACGCTCTGCTCGTTGCTACTTTGGAGCCACTAT 2109
Db 2341 ACTGTCGACCGCTTTGCGCCGCCGACGCTCTGCTCGTTGCTACTTTGGAGCCACTAT 2400
QY 2110 CGACTACGCGATATGCGGACCCACACCCGCTCTGTGGATCTATCGATCTAAATGTAAATG 2169
Db 2401 CGACTACGCGATATGCGGACCCACACCCGCTCTGTGGATCTATCGATCTAAATGTAAATG 2460
QY 2170 TAAATCTCTAAATTAATTAATAGTCCCAAGTTTCTCCATACGAACCTTAAACAGCATTCG 2229
Db 2461 TAAATCTCTAAATTAATTAATAGTCCCAAGTTTCTCCATACGAACCTTAAACAGCATTCG 2520
QY 2230 GGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGTGCGCAATATGTTTCA 2289
Db 2521 GGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGTGCGCAATATGTTTCA 2580
QY 2290 GTCCCTCAGGAGTTACGCTCTTGTGAAGTGTAACTTTCTGGAAGTTTGCAGTGTAACTC 2349
Db 2581 GTCCCTCAGGAGTTACGCTCTTGTGAAGTGTAACTTTCTGGAAGTTTGCAGTGTAACTC 2640
QY 2350 CGCTGTATTGACGGGCATATCCGTACGTTGGCAAGTGTGTTGGTACCGAGAGTAAAT 2409

Db 2641 CGCTGTATTGACGGGCATATCCGTACGTTGGCAAGTGTGTTGGTACCGGAGAGTAAAT 2700
QY 2410 CTCCTCAACTCTCTCGAGAGTAGGACCAACAAACAGATCCAGCGTGTGTACTTTGAT 2469
Db 2701 CTCCTCAACTCTCTCGAGAGTAGGACCAACAAACAGATCCAGCGTGTGTACTTTGAT 2760
QY 2470 CAACATAAGAGAAGCATTTCTCGATTTGAGATCAAGTGTTCAGGAGCGTACTGATTTGG 2529
Db 2761 CAACATAAGAGAAGCATTTCTCGATTTGAGAGTCAAGTGTTCAGGAGCGTACTGATTTGG 2820
QY 2530 ACATTTCCAAAGCCTGCTGTTAGGTTGCAACCGATAGGTTGTAGAGTGTGCAATACACT 2589
Db 2821 ACATTTCCAAAGCCTGCTGTTAGGTTGCAACCGATAGGTTGTAGAGTGTGCAATACACT 2880
QY 2590 TCGGTACATTTTCAACCCCTTGGCAACTGCAAGCTTTGTTGTGAACAGCATCTTCAATTC 2649
Db 2881 TCGGTACATTTTCAACCCCTTGGCAACTGCAAGCTTTGTTGTGAACAGCATCTTCAATTC 2940
QY 2650 TGGCAAGCTCTCTTGTCTGTCTATATCGACAGCCAAACAGAAATCACCTGGGAATCAATACCAT 2709
Db 2941 TGGCAAGCTCTCTTGTCTGTCTATATCGACAGCCAAACAGAAATCACCTGGGAATCAATACCAT 3000
QY 2710 GTTCAGCTTGAGACAGAGAGTCTGAGGCAACGAAATCTGGATCAGCGTATTTATCAGCAA 2769
Db 3001 GTTCAGCTTGAG - CAGAAAGGTTGAGGCAACGAAATCTGGATCAGCGTATTTATCAGCAA 3059
QY 2770 TAACTAGAACTTCAGAAAGCCCGCAGCAGCATGTCTAATACTACACAGGCGCTCATGTGTAT 2829
Db 3060 TAACTAGAACTTCAGAAAGCCCGCAGCAGCATGTCTAATACTACACAGGCGCTCATGTGTAT 3119
QY 2830 TTTGAACCATCATCTTTGGCAGCAGTAAACGAACTGGTTTCTTGGACCAAAATTTTGTCTAC 2889
Db 3120 TTTGAACCATCATCTTTGGCAGCAGTAAACGAACTGGTTTCTTGGACCAAAATTTTGTCTAC 3179
QY 2890 ACTTAGGAACAGTTTCTGTTCCGTAAGCCATAGCAGCTACTGCTTGGCGGCTCTCTGTCTA 2949
Db 3180 ACTTAGGAACAGTTTCTGTTCCGTAAGCCATAGCAGCTACTGCTTGGCGGCTCTCTGTCTA 3239
QY 2950 GCACGATACACTTAGCACAACCTTTGTTGGGCAACGATAGATCACTTCTGGGGTAAAGGTAC 3009
Db 3240 GCACGATACACTTAGCACAACCTTTGTTGGGCAACGATAGATCACTTCTGGGGTAAAGGTAC 3299
QY 3010 CATCTCTTTAGGTGAGATGCAAAAACAAATTTCTTTGCAACCCAGCACTTTTGCAGGAA 3069
Db 3300 CATCTCTTTAGGTGAGATGCAAAAACAAATTTCTTTGCAACCCAGCACTTTTGCAGGAA 3359
QY 3070 CACCAGCATCAGGGAAGTGAAGGCAAGATTTGCGGTTTCCACAGGAATATAGAGGCCAA 3129
Db 3360 CACCAGCATCAGGGAAGTGAAGGCAAGATTTGCGGTTTCCACAGGAATATAGAGGCCAA 3419
QY 3130 CTTTCTCAATAGCTTTCCAAAACGAGAGCAGACTACACAGGCGCAAGTCTCACTTTGCA 3189
Db 3420 CTTTCTCAATAGCTTTCCAAAACGAGAGCAGACTACACAGGCGCAAGTCTCACTTTGCA 3479
QY 3190 AGCTCTCGTTAGTTGAGCTTTCATGGAATTTCTGAGCTTATCTATAGAGAGATCAATGG 3249
Db 3480 AGCTCTCGTTAGTTGAGCTTTCATGGAATTTCTGAGCTTATCTATAGAGAGATCAATGG 3539
QY 3250 CTTCTTAAAGCTTATCTGGCAATTTGCAATGTTCTCTGGGAAAAGGAGCTTCTAACACAG 3309
Db 3540 CTTCTTAAAGCTTATCTGGCAATTTGCAATGTTCTCTGGGAAAAGGAGCTTCTAACACAG 3599
QY 3310 GTGTCTTCAAAAGCAGCTCCATCAAACTTTGGCAGTTAGTTTCTTAAAGGCTTTGTCCACAT 3369
Db 3600 GTGTCTTCAAAAGCAGCTCCATCAAACTTTGGCAGTTAGTTTCTTAAAGGCTTTGTCCACAT 3659
QY 3370 TTTGACGAACATTTGTCGCAATTTGGTTTGAATAATCCATAATCTGTTCGGTTTCTGGA 3429
Db 3660 TTTGACGAACATTTGTCGCAATTTGGTTTGAATAATCCATAATCTGTTCGGTTTCTGGA 3719
QY 3430 TAGGACGAGAGGCGCATCTTCAATTTCTGTAGGAGGCGCTTAGAAACGTCATTTTTCG 3489

Db	3720	TAGGACGACGAGGCACTTCAATTTCTTTGTGAGGAGGCTTAGAAACGCTCAATTTTGC	3779	Qy	4569	TGCACATTAACCTTGAAGCTCAGTCGATTTAGTGAACCTTGATCAGGTTCTGCGAGCTGCTCA	4628
Qy	3490	ACAAATTCATACGACCTTCAGAGGAGCTCTCTTTAGTTTGGATTTCTTTCTTTAGTTGTT	3549	Db	4859	TGCACATTAACCTTGAAGCTCAGTCGATTTAGTGAACCTTGATCAGGTTCTGCGAGCTGCTCA	4918
Db	3780	ACAAATTCATACGACCTTCAGAGGAGCTCTCTTTAGTTTGGATTTCTTTCTTTAGTTGTT	3839	Qy	4629	GCAGCATAGGAAACACAGGCTTTTCTTACCAAACTCAAGGAATTAACAACCTCTGCAACA	4688
Qy	3550	CTTTGGTATACCTTGGCTTGGCATCTCTTTCTTTCTTTAGTGAACCTTTAGGACCTTCAAT	3609	Db	4919	GCAGCATAGGAAACACAGGCTTTTCTTACCAAACTCAAGGAATTAACAACCTCTGCAACA	4978
Db	3840	CTTTGGTATACCTTGGCTTGGCATCTCTTTCTTTCTTTAGTGAACCTTTAGGACCTTCAAT	3899	Qy	4689	CTTTGGTATAGCAGTGAAGGAATTCATCTTTGAAGTCGACAGTGAAGTGTAGTCT	4748
Qy	3610	CCAGTTTCTCTCCACCTCGTCCAAACGTCACACGCTACTTGGCACATCTAACTAAAGCAA	3669	Db	4979	CTTTGGTATAGCAGTGAAGGAATTCATCTTTGAAGTCGACAGTGAAGTGTAGTCT	5038
Db	3900	CCAGTTTCTCTCCACCTCGTCCAAACGTCACACGCTACTTGGCACATCTAACTAAAGCAA	3959	Qy	4749	TGAGAAATTTCTCAAGCCGTAATTTTATTCAGTCAGTCAGTCATCAGAGATCTCTTAC	4808
Qy	3670	AATAAATAAGTCAGCAATTCACAGGTATATCTTCTTTGGATTTAGCTTTTGCAGATT	3729	Db	5039	TGAGAAATTTCTCAAGCCGTAATTTTATTCAGTCAGTCAGTCATCAGAGATCTCTTAC	5098
Db	3960	AATAAATAAGTCAGCAATTCACAGGTATATCTTCTTTGGATTTAGCTTTTGCAGATT	4019	Qy	4809	GCAGCATAGGAAACACAGGCTTTTCTTACCAAACTCAAGGAATTAACAACCTCTGCAACA	4868
Qy	3730	CATCAGCTTCTCTCTCTTTAGGTTTACGATTC - AACAAAATCTTCTGCTCAAAATAACCGTTTG	3788	Db	5099	GCAGCATAGGAAACACAGGCTTTTCTTACCAAACTCAAGGAATTAACAACCTCTGCAACA	5158
Db	4020	CATCAGCTTCTCTCTCTTTAGGTTTACGATTC - AACAAAATCTTCTGCTCAAAATAACCGTTTG	4079	Qy	4869	GCAGCATAGGAAACACAGGCTTTTCTTACCAAACTCAAGGAATTAACAACCTCTGCAACA	4928
Qy	3789	GTATAAGAACCTTCTGGAGCATTTGCTCTTACGATCCCAAGGTTGCTTCCATGGCTCTA	3848	Db	5159	GCAGCATAGGAAACACAGGCTTTTCTTACCAAACTCAAGGAATTAACAACCTCTGCAACA	5218
Db	4080	GTATAAGAACCTTCTGGAGCATTTGCTCTTACGATCCCAAGGTTGCTTCCATGGCTCTA	4138	Qy	4929	AGCGCTTGTTCGCGCTGGGTATGTTGTCAGGCCCGCTGCGCGGGAAGTCTGTTGGCGCC	4988
Qy	3849	AGACCTTTGATTTGGCCAAACAGGAAGTGGTTTCCAAAGTCACAGAAACCAACCTGTT	3908	Db	5219	AGCGCTTGTTCGCGCTGGGTATGTTGTCAGGCCCGCTGCGCGGGAAGTCTGTTGGCGCC	5278
Db	4139	AGACCTTTGATTTGGCCAAACAGGAAGTGGTTTCCAAAGTCACAGAAACCAACCTGTT	4198	Qy	4989	ATCTCTCTG	4997
Qy	3909	TGTTCAACCAACAAATTTCAAGCAGTCTCCATCACAATCCAAATTCGATACCCAGCAACTTT	3968	Db	5279	ATCTCTCTGACCTGCGAGGGGGGGGGGAAAGCAACGCTTGTCTCAAAATCTCTGAT	5338
Db	4199	TGTTCAACCAACAAATTTCAAGCAGTCTCCATCACAATCCAAATTCGATACCCAGCAACTTT	4258	Qy	4998	-----	4997
Qy	3969	TGAGTTGCTCAGATGTAGCACTTTATACCAAAACCGTGACGACGAGATTTGTAGACT	4028	Db	5339	GTTCATTTGCACAGATAAATAATATCATCATGAACAATAAACTGTCTGCTTACATAA	5398
Db	4259	TGAGTTGCTCAGATGTAGCACTTTATACCAAAACCGTGACGACGAGATTTGTAGACT	4318	Qy	4998	-----	4997
Qy	4029	CCAGTTTGTCTTATAGCTTCGGGATAGACTTTTGGACGAGTACACAGGCCCAAC	4088	Db	5399	ACAGTAATAACAGGGGTATTATGAGCCATATTCAACGGGAAACGCTTCTCTCAAGGCCG	5458
Db	4319	CCAGTTTGTCTTATAGCTTCGGGATAGACTTTTGGACGAGTACACAGGCCCAAC	4378	Qy	4998	-----	4997
Qy	4089	GAGTAATTTAGAGAGTTCAGCCACCAAGTAGTAATAGAACCATCGGGCGGTCTAGTAGTC	4148	Db	5459	GATTAATTTCAACATGGATGCTGATTTATATGGGTATTAATGGGTCTCGGCTCGGATATGTCG	5518
Db	4379	GAGTAATTTAGAGAGTTCAGCCACCAAGTAGTAATAGAACCATCGGGCGGTCTAGTAGTC	4438	Qy	4998	-----	4997
Qy	4149	AAAGCGCCCAACAAATTTCACTGACGAGGAACTTTTGGATCTTTCAGAAAGTTCGTAT	4208	Db	5519	GGCAATCAGGTGCGACAACTCTATCGATTTGATGGGAAGCCCGATGCGCCAGAGTTGTTTC	5578
Db	4439	AAAGCGCCCAACAAATTTCACTGACGAGGAACTTTTGGATCTTTCAGAAAGTTCGTAT	4498	Qy	4998	-----	4997
Qy	4209	TCAGTAGTCAATTTGCCGAGCATCAATATGGGATTTATACCAGAAACAGTGGAAAGTC	4268	Db	5579	TGAAACATGGCAAAAGTAGCGTTGCCCATCAAGCATTTTATTCGTAATGATGATG	5638
Db	4499	TCAGTAGTCAATTTGCCGAGCATCAATATGGGATTTATACCAGAAACAGTGGAAAGTC	4558	Qy	4998	-----	4997
Qy	4269	ACATCTACCACTTTGGCGTCTCAGAAAGAGCATAAACGTTCTACTACCGCCATTAGTG	4328	Db	5639	GGCTGACGGAATTTATGCTCTTCCGACCATCAAGCATTTTATTCGTAATGATGATG	5698
Db	4559	ACATCTACCACTTTGGCGTCTCAGAAAGAGCATAAACGTTCTACTACCGCCATTAGTG	4618	Qy	4998	-----	4997
Qy	4329	AAACTTTTCAATCGCCAGTGGAGAAAGAAAGCAACAGCATTTACTAGCATTTAGCGGC	4388	Db	5699	CATGTTTACTCACCCTGCGATCCCGGGGAAAACAGCATTTCCAGGTATTAGAAATATC	5758
Db	4619	AAACTTTTCAATCGCCAGTGGAGAAAGAAAGCAACAGCATTTACTAGCATTTAGCGGC	4678	Qy	4998	-----	4997
Qy	4389	AAAGGATGCAACTTTATCAACACGAGGTCTTATAGATAAACCTTAGCGCTGGGATCATCTT	4448	Db	5759	CTGATTCAGGTGAATAATTTGTTGATGCGCTGCGAGTGTCTCTGCGCGGTTGCATTGCA	5818
Db	4679	AAAGGATGCAACTTTATCAACACGAGGTCTTATAGATAAACCTTAGCGCTGGGATCATCTT	4738	Qy	4998	-----	4997
Qy	4449	TGGACAACTTTCTGCGCAAAATCTAGGTCGCAAAATCACTTCAATTCATACCATTTTGTAC	4508	Db	5819	TTCTGTTTGTAAATTTGCTTTTAAACAGCGATCGCGTATTTCTGCTCTGCTCAGCGCAAT	5878
Db	4739	TGGACAACTTTCTGCGCAAAATCTAGGTCGCAAAATCACTTCAATTCATACCATTTTGTAC	4798	Qy	4998	-----	4997
Qy	4509	AACTTGAGCAAGTTGTCGATCAGCTCTCTCAAAATTTGCTCTCTGTAACCGATGACTCAACT	4568	Db	5879	CACGAATGAATAACGGTTTGGTTGATGCGGAGTGAATTTGATGACGAGCGTAAATGGCTGGC	5938
Db	4799	AACTTGAGCAAGTTGTCGATCAGCTCTCTCAAAATTTGCTCTCTGTAACCGATGACTCAACT	4858				

8099 AGAACCOCCTTTCAGCCGACCGCTCGCCCTTATCCGGTAACATATCGTCTTGAAGTCCA 8158
6629 ACCCGGTAAAGACACAGACTTATCGCCACTCGGACAGACCCACTGGTAAACAGGATTAGCAGAG 6688
8159 ACCCGGTAAAGACACAGACTTATCGCCACTCGGACAGACCCACTGGTAAACAGGATTAGCAGAG 8218
6689 CGAGTATGTAGCGGCTGCTACAGAGTCTTGAAGTGTGTCCTAACTACCGCTACACTA 6748
8219 CGAGTATGTAGCGGCTGCTACAGAGTCTTGAAGTGTGTCCTAACTACCGCTACACTA 8278
6749 GAAGACAGTATTTGGTATCTCGCTCTGCTGAGCCAGTATACCTTCGGAAAAAGAGTTG 6808
8279 GAAGACAGTATTTGGTATCTCGCTCTGCTGAGCCAGTATACCTTCGGAAAAAGAGTTG 8338
6809 CTAGCTCTTGATTCGGGCAAAACCAACCGCTGTAGCGGTGGTTTTTTTCTTTGCAAGC 6868
8339 CTAGCTCTTGATTCGGGCAAAACCAACCGCTGTAGCGGTGGTTTTTTTCTTTGCAAGC 8398
6869 ACCAGATTACCGGCAAAAGAGATCTCAGAGATCCCTTGATCTTTCTACGGGT 6928
8399 ACCAGATTACCGGCAAAAGAGATCTCAGAGATCCCTTGATCTTTCTACGGGT 8458
6929 CTGACGCTCAGTGAAGCAAACTCAGCTTAAGGATTTTGGTCATGAGATTAACAATA 6988
8459 CTGACGCTCAGTGAAGCAAACTCAGCTTAAGGATTTTGGTCATGAGATTAACAATA 8518
6989 GGATCTTCACTAGATTCCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATAT 7048
8519 GGATCTTCACTAGATTCCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATAT 8578
7049 ATGAGTAACTTGGTCTGACAGTACCAATGCTTAATCAGTGAGGACCTATCTCAGCGA 7108
8579 ATGAGTAACTTGGTCTGACAGTACCAATGCTTAATCAGTGAGGACCTATCTCAGCGA 8638
7109 TCTGTCTATTTCTGTTTCACTCCATAGTGTGCTGACTCCCGTGTGCTGATGATTA 7168
8639 TCTGTCTATTTCTGTTTCACTCCATAGTGTGCTGACTCCCGTGTGCTGATGATTA 8698
7169 GGGGGGCTTACATCTGCCCCAGTGTGCAATCATACCCGAGACCCAGCTCAGCGG 7228
8699 GGGGGGCTTACATCTGCCCCAGTGTGCAATCATACCCGAGACCCAGCTCAGCGG 8758
7229 CTCACAGATTTATCAGCAATAAACCAGCCAGCCGAGGCGGAGCGAGAGTGTCTCTG 7288
8759 CTCACAGATTTATCAGCAATAAACCAGCCAGCCGAGGCGGAGCGAGAGTGTCTCTG 8818
7289 CAACCTTATCCGCTCCATCCAGTCTAATTAATTTGTCGGGAGCTAGAGTAAGTAGTT 7348
8819 CAACCTTATCCGCTCCATCCAGTCTAATTAATTTGTCGGGAGCTAGAGTAAGTAGTT 8878
7349 CGCCAGTAAATAGTTTGGCAACGTTGTTGCCATTGCTGAGCATCTGTTGTCACGCT 7408
8879 CGCCAGTAAATAGTTTGGCAACGTTGTTGCCATTGCTGAGCATCTGTTGTCACGCT 8938
7409 CGTGTGTTGGTATGGCTTCATTCAGTCCGGTTCCTCAACAGATCAAGGCGAGTTACATAT 7468
8939 CGTGTGTTGGTATGGCTTCATTCAGTCCGGTTCCTCAACAGATCAAGGCGAGTTACATAT 8998
7469 CCCCCATGTTGCAAAAAGCGGTTAGCTCTCGCTCCCTCCGATCGTTGTCAGAGTA 7528
8999 CCCCCATGTTGCAAAAAGCGGTTAGCTCTCGCTCCCTCCGATCGTTGTCAGAGTA 9058
7529 AGTGGCCGAGTGTATCACTCATGTTTATGGCAGCATGATCAATTTCTTACTGTCA 7588
9059 AGTGGCCGAGTGTATCACTCATGTTTATGGCAGCATGATCAATTTCTTACTGTCA 9118
7589 TGGCATCCGTAGATGCTTTCTGTGACTGTGGTGTAGTACTCAACCAAGTCAATTCAGAT 7648
9119 TGGCATCCGTAGATGCTTTCTGTGACTGTGGTGTAGTACTCAACCAAGTCAATTCAGAT 9178
7649 AGTGTATGGCGGACCGAGTGTGCTTTGCCCGGCTCAACCGGATTAATACCGGCCAC 7708
9179 AGTGTATGGCGGACCGAGTGTGCTTTGCCCGGCTCAACCGGATTAATACCGGCCAC 9238

7709 ATAGCAGAACTTTAAAGTGCTCATCATTTGGAAAACTTTCTTGGGGCGAAAACTCTCAA 7768
9239 ATAGCAGAACTTTAAAGTGCTCATCATTTGGAAAACTTTCTTGGGGCGAAAACTCTCAA 9298
7769 GATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGACCCCACTGATCTT 7828
9299 GATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGACCCCACTGATCTT 9358
7829 CAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAAAGGCAAAATGCCG 7888
9359 CAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAAAGGCAAAATGCCG 9418
7889 CAAAAAGGGAATAAGGGGACACGGAATTTGAATCTATCATCTCTTCTTTTCAAT 7948
9419 CAAAAAGGGAATAAGGGGACACGGAATTTGAATCTATCATCTCTTCTTTTCAAT 9478
7949 ATTATTGAACATTTATCAGGGTTATTTCTCATGAGCGGATACATATTGATGTATTT 8008
9479 ATTATTGAACATTTATCAGGGTTATTTCTCATGAGCGGATACATATTGATGTATTT 9538
8009 AGAAAAATAAACAATAGGGTTTCCGGGACATTTCCCGAAAAAGTGCCACCTGAGTCT 8068
9539 AGAAAAATAAACAATAGGGTTTCCGGGACATTTCCCGAAAAAGTGCCACCTGAGTCT 9598
8069 AAGAAACCATTTATCATGACATTAACCTTATAAAATAGGCGTATCAGAGCCCTTTC 8128
9599 AAGAAACCATTTATCATGACATTAACCTTATAAAATAGGCGTATCAGAGCCCTTTC 9658
8129 GTCTTCAAGAAATTAATCTCATGTTTGACAGCTTATCATGTAAGCTGATCATGTGG 8188
9659 GTCTTCAAGAAATTAATCTCATGTTTGACAGCTTATCATGTAAGCTGATCATGTGG 9718
8189 TATTGTGAATAGACCGGATCGGACACTGGAATAACAGTATTAATTCG 8241
9719 TATTGTGAATAGACCGGATCGGACACTGGAATAACAGTATTAATTCG 9771

RESULT 10

AX417106

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

i. 9776

/organism="unidentified"

/mol_type="unassigned DNA"

/db_xref="taxon:32644"

/note="Plasmid p55"

ORIGIN

Query Match

Best Local Similarity

Matches 8101; Conservative

2

1

62

61

121

120

77.5%; Score 6383.8; DB 6; Length 9776;

82.9%; Pred. No. 0;

0; Mismatches 137; Indels 1535; Gaps 10;

GATCTAATCCAAAGACGAAAGGTTGAATGAACCTTTTGGCATCCGACATCCACAGG 61

GATCTAATCCAAAGACGAAAGGTTGAATGAACCTTTTGGCATCCGACATCCACAGG 60

TCCATTCCTCACATAGTGCCTCAACGCAACAGAGGGGATACACTAGCAGCAGCGTT 121

TCCATTCCTCACATAGTGCCTCAACGCAACAGAGGGGATACACTAGCAGCAGCGTT 120

9776 bp DNA linear PAT 14-JUN-2002

Sequence 31 from Patent WO0198344.

AX417106

AX417106.1 GI:21449693

unidentified

unidentified

unclassified.

Ling, L.E. and Sanicola-Nadel, M.

Angiogenesis-modulating compositions and uses

Patent: WO 0198344-A 31 27-DEC-2001;

BIOGEN, INC. (US)

Location/Qualifiers

i. 9776

/organism="unidentified"

/mol_type="unassigned DNA"

/db_xref="taxon:32644"

/note="Plasmid p55"

122 GCAACGAGGACCTCCACTCTCTCTCTCTCAACACCCACTTTTGCCATCGAAAAACCA 181
121 GCAACGAGGACCTCCACTCTCTCTCTCAACACCCACTTTTGCCATCGAAAAACCA 180
182 GCCAGATTATGGGCTTGATGGAGCTGCTCATTCGAATCTCTTCTATTAGGCTACTAA 241
181 GCCAGATTATGGGCTTGATGGAGCTGCTCATTCGAATCTCTTCTATTAGGCTACTAA 240
242 CACCATGACTTTATTTAGGCTCTCTATCTGCGCCCTCTGGCGAGGTTCAATGTTGTTAT 301
241 CACCATGACTTTATTTAGGCTCTCTATCTGCGCCCTCTGGCGAGGTTCAATGTTGTTAT 300
302 TTCGAATGCAACAGCTCCGCATTACACCCGAACATCACTCCAGATGAGGGCTTCTGA 361
301 TTCGAATGCAACAGCTCCGCATTACACCCGAACATCACTCCAGATGAGGGCTTCTGA 360
362 GTGTGGGCTCAAAATAGTTTCATGTTCCCAATGCGCCAAACTGACAGTTTAAACGCTG 421
361 GTGTGGGCTCAAAATAGTTTCATGTTCCCAATGCGCCAAACTGACAGTTTAAACGCTG 420
422 TCTTGGAACTTAATATGACAAAGCTGATCTCATCCAAGATGAATTAAGTTTGGTTGCT 481
421 TCTTGGAACTTAATATGACAAAGCTGATCTCATCCAAGATGAATTAAGTTTGGTTGCT 480
482 TGAATGCTAAGCGGCTGCTGTCARAAAGAACTTCARAAAGTCGCATACGTTGTC 541
481 TGAATGCTAAGCGGCTGCTGTCARAAAGAACTTCARAAAGTCGCATACGTTGTC 540
542 TTGTTTGTATTTGATGACGAATGCTCAAAAATATCTCAATTAATGTTAGCGAGTCTC 601
541 TTGTTTGTATTTGATGACGAATGCTCAAAAATATCTCAATTAATGTTAGCGAGTCTC 600
602 TCTATCGCTTCTGAACCCCGTGTGACCTGCGGAACCGCAATGGGGAAACACCCGCTT 661
601 TCTATCGCTTCTGAACCCCGTGTGACCTGCGGAACCGCAATGGGGAAACACCCGCTT 660
662 TTGTGATGATTTGATGATGCTCCACATTTGATGCTTCCAAAGATCTGTGGTGGGAATACTG 721
661 TTGTGATGATTTGATGATGCTCCACATTTGATGCTTCCAAAGATCTGTGGTGGGAATACTG 720
722 CTGATAGCTTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTGACAGCAATA 781
721 CTGATAGCTTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTGACAGCAATA 780
782 TATAACAGAGGAGCTGCCCTGCTCTTAAACCTTTTTTTTATCATCATTTATGCTTA 841
781 TATAACAGAGGAGCTGCCCTGCTCTTAAACCTTTTTTTTATCATCATTTATGCTTA 840
842 CTTTCATAATTCGCACTGGTTCCAAATGACAAAGCTTTTGATTTTAAACGACTTTTAAACGAC 901
841 CTTTCATAATTCGCACTGGTTCCAAATGACAAAGCTTTTGATTTTAAACGACTTTTAAACGAC 900
902 AACTTGAGAGATCAAAAACAACTAATTTATGCAAGATCCAAACGATGAGATTTCCTT 961
901 AACTTGAGAGATCAAAAACAACTAATTTATGCAAGATCCAAACGATGAGATTTCCTT 960
962 CAATTTTACTGCACTGTTTATTCGAGCATCTCCGCAATAGCTGCTCCAGTCAACACTA 1021
961 CAATTTTACTGCACTGTTTATTCGAGCATCTCCGCAATAGCTGCTCCAGTCAACACTA 1020
1022 CAACAGAGATGAACGGCACAATTCGGCTGAAGCTGCTCATCGGTTACTCAGATTAG 1081
1021 CAACAGAGATGAACGGCACAATTCGGCTGAAGCTGCTCATCGGTTACTCAGATTAG 1080
1082 AAGGGGATTTGATGCTGCTGTTTGGCAATTTCCACAGACACAATAAAGGGTTATTTGT 1141
1081 AAGGGGATTTGATGCTGCTGTTTGGCAATTTCCACAGACACAATAAAGGGTTATTTGT 1140
1142 TTATAAATACTACTATTCGAGCATTTGCTGTAAAGAAAGGGGTATCTCTCGAGAAA 1201
1141 TTATAAATACTACTATTCGAGCATTTGCTGTAAAGAAAGGGGTATCTCTCGAGAAA 1200
1202 GAGAGGCTGAAGCCCGG----- 1219

1201 GATCGGACCGGCGAGGGGTTCCGGAAAGAGAGGACACCCCAAAAGTGACCCCTTTAG 1260
1220 -----CCCGAGGCCAGGCCCGCCAG 1242
1261 CCTACAAGCAGTTTATCCCAATGTGGCCGAGAGACCTTAGGCGCCAGCGGAAGGTATG 1320
1243 GAAGAAACAT----- 1252
1321 AAGGGAAGATCTCCAGAAACTCCGAGCGATTTAAGGAATCAACCCCAATTACAACCCCG 1380
1253 -----GTATCATGCGCCCGCGAGAGGATAAATTGTGTGGCTTCCCGGTGTC 1299
1381 ACATCATATTTAAGGATGAAGAAACACCGGAGCGGACAGCTGATGACTCAGAGGTGA 1440
1300 AC-----CGCCAGCGATGACGAGAGAGGTTGCTGTTT 1334
1441 AGGACAAGTTGAACGCTTTTGGCCATCTCGGTGATGAACCGAGTGGCCAGGAGTGAACCTGC 1500
1335 TGAATGACATGTCGCGGATTTCCCGTGG----- 1362
1501 GGGTACCGAGGGCTGGGACGAAGATGGCCACCACTCAGAGGAGTCTCTGCATACGAGG 1560
1363 ----- 1362
1561 GCGCGCAGTGGACATCAACAGCTGTGACCGCGACCGCAAGTACGGCATGCTGGGCC 1620
1363 -----TGCTTCCACCCCATGGCCATCGAGAACACTCAAGAAAGAGAT 1405
1621 GCGTGGCGTGGAGCGCGCTTCGACTCGGTGTACTACGATCCAAAGCAATATFCCA 1680
1406 GTCCCTCTTAACATA-----GTGGGTAGAATTCCTTAGG-----CGCGCCGCGA 1449
1681 GCTCGGTGAACGACAGAACTCGGTGGCGGCCAAATCGGAGGCTGATTCGCGCCGCGA 1740
1450 ATTAATTCGCTTAGACATGACTGTTCCCTCAGTTCAAGTTGGGCACTTACAGAGAACCG 1509
1741 ATTAATTCGCTTAGACATGACTGTTCCCTCAGTTCAAGTTGGGCACTTACAGAGAACCG 1800
1510 GTCTTGTAGATTTCAATCAAGAGATGTCAGAAATGCCATTTGCTGAGAGATGCAGGCT 1569
1801 GTCTTGTAGATTTCAATCAAGAGATGTCAGAAATGCCATTTGCTGAGAGATGCAGGCT 1860
1570 TCATTTTGTATCTTTTATTTGTAACCTATATAGTATAGATTTTTTTTTTGTCTATTTG 1629
1861 TCATTTTGTATCTTTTATTTGTAACCTATATAGTATAGATTTTTTTTTTGTCTATTTG 1920
1630 TTTCTTCTGTHAGAGCTTGTCTGATCAGCCTATCTCGAGCTGATGATATCTCTG 1689
1921 TTTCTTCTGTHAGAGCTTGTCTGATCAGCCTATCTCGAGCTGATGATATCTCTG 1980
1690 GTAGGGTTTGGGAAAATCAITTCGAGTTTGTATGTTTTTCTTGGTATTTTCCCACTCTCT 1749
1981 GTAGGGTTTGGGAAAATCAITTCGAGTTTGTATGTTTTTCTTGGTATTTTCCCACTCTCT 2040
1750 CAGAGTACAGAGATTAAGTGAAGTTGTTTGTGCAAGCTTATCGATAGCTTTAATG 1809
2041 CAGAGTACAGAGATTAAGTGAAGTTGTTTGTGCAAGCTTATCGATAGCTTTAATG 2100
1810 CGGTAGTTTATCACAGTTAAATTTGCTAACCGAGTCAGGCACCGTGTATGAAATCTAACAA 1869
2101 CGGTAGTTTATCACAGTTAAATTTGCTAACCGAGTCAGGCACCGTGTATGAAATCTAACAA 2160
1870 TGGCTCATCTGCTGCGCACCGCTCACTCGATGCTGATGAGCTAGCTTGGTTAT 1929
2161 TGGCTCATCTGCTGCGCACCGCTCACTCGATGCTGATGAGCTAGCTTGGTTAT 2220
1930 GCGGTACTGCGCGGCTCTTTCGGGGATATCTGTCATTTCCGACAGCATCGCCAGTCACTA 1989
2221 GCGGTACTGCGCGGCTCTTTCGGGGATATCTGTCATTTCCGACAGCATCGCCAGTCACTA 2280
1990 TGGCGTGTCTGAGCGCTATATGCGTTTGTATGCAATTTCTATGCGCACCGCTTCTCGGAGC 2049

2281 TGCGGTGCTGCTAGCGCTATATGCGTGTGATGCAATTTCTATGCGCACCGGTTCTCGGAGC 2340
2050 ACTGTCCGACCGCTTTGGCGCGCCGAGTCTGTGCTCGCTTCTGCTTCTGAGCGCACTAT 2109
2341 ACTGTCCGACCGCTTTGGCGCGCCGAGTCTGTGCTCGCTTCTGCTTCTGAGCGCACTAT 2400
2110 CGACTACCGGATCATGGCGACACACCCGCTCTGTGGATCTATCGAATCTAAATGTAAAT 2169
2401 CGACTACCGGATCATGGCGACACACCCGCTCTGTGGATCTATCGAATCTAAATGTAAAT 2460
2170 TAAATCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2229
2461 TAAATCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2520
2230 GGTGAGCATCTAGACCTTTCAACAGCAGCGAGATCCATCACTGCTTGGCCAAATGTTTCA 2289
2521 GGTGAGCATCTAGACCTTTCAACAGCAGCGAGATCCATCACTGCTTGGCCAAATGTTTCA 2580
2290 GTCCCTCAGGAGTTACGCTTGTGAGTGTGAACTTCTGGAAGTTGCACTGTTAACTC 2349
2581 GTCCCTCAGGAGTTACGCTTGTGAGTGTGAACTTCTGGAAGTTGCACTGTTAACTC 2640
2350 CGCTGTATTGACGGGCATATCCGTACGTTGGCAAGTGTGGTACCGGAGGAGTAAT 2409
2641 CGCTGTATTGACGGGCATATCCGTACGTTGGCAAGTGTGGTACCGGAGGAGTAAT 2700
2410 TTCCCAACTCTCTGGAGTAGGACCAACAAACACAGATCCAGCGGTGTTGTTGAT 2469
2701 CTCCCAACTCTCTGGAGTAGGACCAACAAACACAGATCCAGCGGTGTTGTTGAT 2760
2470 CAACATAGAGAGCAATCTCGATTTTCAGAGTCAAGTGTTCAGGAGCGTACTGATTGG 2529
2761 CAACATAGAGAGCAATCTCGATTTTCAGAGTCAAGTGTTCAGGAGCGTACTGATTGG 2820
2530 ACATTTCAAAGCCTGCTCGTAGTTCGAAACCGATAGGTTGTAGAGTGTGCAATCACT 2589
2821 ACATTTCAAAGCCTGCTCGTAGTTCGAAACCGATAGGTTGTAGAGTGTGCAATCACT 2880
2590 TCGGTACAAATTCACCTTTGGCACTGACAGCTTGGTGTGACAGCATCTTCAATTC 2649
2881 TCGGTACAAATTCACCTTTGGCACTGACAGCTTGGTGTGACAGCATCTTCAATTC 2940
2650 TGGCAAGCTCCTTGTCTGTATATCGACAGCCAAACAGAACTACCTGGGAATCAATACCAT 2709
2941 TGGCAAGCTCCTTGTCTGTATATCGACAGCCAAACAGAACTACCTGGGAATCAATACCAT 3000
2710 GTTCAGTTGAGACAGAGGTTGAGGCAACGAAATCTGGATCAGCGTATTTATCAGCAA 2769
3001 GTTCAGTTGAGACAGAGGTTGAGGCAACGAAATCTGGATCAGCGTATTTATCAGCAA 3059
2770 TAACTAGAACTTCAGAGGCCAGAGCATGTCAATCTACTACAGAGGCTCATGTTCAT 2829
3060 TAACTAGAACTTCAGAGGCCAGAGCATGTCAATCTACTACAGAGGCTCATGTTCAT 3119
2830 TTTGAAACCATCATCTTGGCAGCAGTAAACGAACTGGTTTCTGGACCAATATTTGTAC 2889
3120 TTTGAAACCATCATCTTGGCAGCAGTAAACGAACTGGTTTCTGGACCAATATTTGTAC 3179
2890 ACTTAGGACAGTTTCTGTTCCGTAAGCATAGCACTACTGCTGGGGGCTCTGCTA 2949
3180 ACTTAGGACAGTTTCTGTTCCGTAAGCATAGCACTACTGCTGGGGGCTCTGCTA 3239
2950 GCACGATACACTTAGCACCACCTTTGGGCAACGATAGATGACTTCTGGGGTAAAGGTAC 3009
3240 GCACGATACACTTAGCACCACCTTTGGGCAACGATAGATGACTTCTGGGGTAAAGGTAC 3299
3010 CATCTCTTAGGTGAGTGAACCAATTTCTTTGGACCGACCACTTTGGCAGGAA 3069
3300 CATCTCTTAGGTGAGTGAACCAATTTCTTTGGACCGACCACTTTGGCAGGAA 3359
3070 CACCAGCATCAGGGAAGTGAAGCGAATTTGGGTTCCACCGAATATAGAGGCCAA 3129
3360 CACCAGCATCAGGGAAGTGAAGCGAATTTGGGTTCCACCGAATATAGAGGCCAA 3419

3130 CTTTCTCAATAGTCTTTGCAAAACAGAGAGCAGACTACACAGGGCAAGTCTCAACTTGA 3189
3420 CTTTCTCAATAGTCTTTGCAAAACAGAGAGCAGACTACACAGGGCAAGTCTCAACTTGA 3479
3190 ACCTCTCCGTTAGTTCAGCTTCAATGAATTTCTGAGTTCATCTATAGAGAGTCAATGG 3249
3480 ACCTCTCCGTTAGTTCAGCTTCAATGAATTTCTGAGTTCATCTATAGAGAGTCAATGG 3539
3250 CTCTCTTAAGCTTATCTGGCAATTTGCAATAGTTCCTCTGGAAAGAGCTTCTAAACACAG 3309
3540 CTCTCTTAAGCTTATCTGGCAATTTGCAATAGTTCCTCTGGAAAGAGCTTCTAAACACAG 3599
3310 GTCTCTTAAGAGCAGCTCCATCAAACTTGGAGTTAGTTCTAAAGGGCTTTGTACCAAT 3369
3600 GTCTCTTAAGAGCAGCTCCATCAAACTTGGAGTTAGTTCTAAAGGGCTTTGTACCAAT 3659
3370 TTTGAGCAATCTGTCAGCAATTTGCAATATTTCCATATCTCTCCGTTTCTGGA 3429
3660 TTTGAGCAATCTGTCAGCAATTTGCAATATTTCCATATCTCTCCGTTTCTGGA 3719
3430 TAGGAGCAAGAGGCGATCTTCAATTTCTTGTGAGGAGGCTTTAGAAACGTCATTTTGC 3489
3720 TAGGAGCAAGAGGCGATCTTCAATTTCTTGTGAGGAGGCTTTAGAAACGTCATTTTGC 3779
3490 ACAATTTCAATACGACTTCAGAGGAGCTTCTTAGTTCGATTTGATTTCTTTAGGTTGT 3549
3780 ACAATTTCAATACGACTTCAGAGGAGCTTCTTAGTTCGATTTGATTTCTTTAGGTTGT 3839
3550 CTTTGTGTATCTCTGGCTTGGCATCTCTTCTCTAGTGAGCTTTAGGAGCTTCAATAT 3609
3840 CTTTGTGTATCTCTGGCTTGGCATCTCTTCTCTAGTGAGCTTTAGGAGCTTCAATAT 3899
3610 CCAGGTTTCTCTCCACCTGTCACCGTCAACGCTACTTGGCACATCTAACTAATGCAA 3669
3900 CCAGGTTTCTCTCCACCTGTCACCGTCAACGCTACTTGGCACATCTAACTAATGCAA 3959
3670 AATAAATAAGTCAGACACATTTCCAGCTATATCTTCTTGGATTTAGCTTCTGCAAGTT 3729
3960 AATAAATAAGTCAGACACATTTCCAGCTATATCTTCTTGGATTTAGCTTCTGCAAGTT 4019
3730 CATCAGCTTCTCTCCATTTTAGCGTTC-AAACAAACTTCGTGCAAAATACCGTTTG 3788
4020 CATCAGCTTCTCTCCATTTTAGCGTTC-AAACAAACTTCGTGCAAAATACCGTTTG 4079
3789 GTATAGAACTTCTGAGGANTGCTTTTACGATCCCAAGGTGCTTCCATGGCTCTA 3848
4080 GTATAGAACTTCTGAGGANTGCTTTTACGATCCCAAGGTGCTTCCATGGCTCTA 4138
3849 AGACCTTTTGAATGGCCAAACAGGAAGTTCGCTTCCAAAGTGACAGAAACCAACCTGTT 3908
4139 AGACCTTTTGAATGGCCAAACAGGAAGTTCGCTTCCAAAGTGACAGAAACCAACCTGTT 4198
3909 TGTTCACCAAAATTTCAAGCAGTCTCCATCAATCCAAATCCGATCCAGCAACTTT 3968
4199 TGTTCACCAAAATTTCAAGCAGTCTCCATCAATCCAAATCCGATCCAGCAACTTT 4258
3969 TGAGTTTCTCCAGATCTGACCTTTTATACCAAAACCGTGACGAGATTTGGTAGACT 4028
4259 TGAGTTTCTCCAGATCTGACCTTTTATACCAAAACCGTGACGAGATTTGGTAGACT 4318
4029 CCAGTTTGTCTTATAGCCTCCGGAATAGACTTTTGGACGAGTACACAGGCCCAAC 4088
4319 CCAGTTTGTCTTATAGCCTCCGGAATAGACTTTTGGACGAGTACACAGGCCCAAC 4378
4089 GAGTAAATTAAGAGGTCAGCCCAAGTAGTGAATAGACCATCGGGCGGTCAAGTAGTC 4148
4379 GAGTAAATTAAGAGGTCAGCCCAAGTAGTGAATAGACCATCGGGCGGTCAAGTAGTC 4438
4149 AAGAGCGCCAAACAAATTTCACTGACAGGGAACCTTTTTCACATCTTTAGAAAGTTCTGAT 4208
4439 AAGAGCGCCAAACAAATTTCACTGACAGGGAACCTTTTTCACATCTTTTTCAGAAAGTTCTGAT 4498

4209 TCAGTAGTCAATTCGCCAGCATCAATTAATGGGATATATACAGAAGCAACAGTGGAGTC 4268
Db
4499 TCAGTAGTCAATTCGCCAGCATCAATTAATGGGATATATACAGAAGCAACAGTGGAGTC 4558
Qy
4269 ACATCTACCAACTTTGGGGTCTCAGAAAAAGCATAAACAGTTCCTACTACCGCCATTAGTG 4328
Db
4559 ACATCTACCAACTTTGGGGTCTCAGAAAAAGCATAAACAGTTCCTACTACCGCCATTAGTG 4618
Qy
4329 AAACCTTTTCAAATCGCCAGTGGAGAGAAAAAGGCACAGCATCTAGCATTAGCGGC 4388
Db
4619 AAACCTTTTCAAATCGCCAGTGGAGAGAAAAAGGCACAGCATCTAGCATTAGCGGC 4678
Qy
4389 AAGGATGCAACTTTATCAACACAGGGTCTATAGATAAACCTAGCGCTGGGATCATCTT 4448
Db
4679 AAGGATGCAACTTTATCAACACAGGGTCTATAGATAAACCTAGCGCTGGGATCATCTT 4738
Qy
4449 TGGACAACTCTTCTGCCAAATCTAGTCCAAATCACTTCAATGTATACCATTTATGTAC 4508
Db
4739 TGGACAACTCTTCTGCCAAATCTAGTCCAAATCACTTCAATGTATACCATTTATGTAC 4798
Qy
4509 AACTTGACAAAGTTGTGATCAGCTCTCTCAAATTTGGTCTCTGTAAACGGATGACTCAACT 4568
Db
4799 AACTTGACAAAGTTGTGATCAGCTCTCTCAAATTTGGTCTCTGTAAACGGATGACTCAACT 4858
Qy
4569 TGCATTAATCTGAAGTCTAGTCTGATGAGTGAACCTGATCAGTGTGTGCAGCTGTCA 4628
Db
4859 TGCATTAATCTGAAGTCTAGTCTGATGAGTGAACCTGATCAGTGTGTGCAGCTGTCA 4918
Qy
4629 GCAGATAGGGAACACCGCTTTTCTCAACAACTCAAGGAATATCAAACTCTGCAACA 4688
Db
4919 GCAGATAGGGAACACCGCTTTTCTCAACAACTCAAGGAATATCAAACTCTGCAACA 4978
Qy
4689 CTTGCGTATGAGGTAGCAAGGAAATGTCTATCTGAAGTCGACAGTGAAGTGTCT 4748
Db
4979 CTTGCGTATGAGGTAGCAAGGAAATGTCTATCTGAAGTCGACAGTGAAGTGTCT 5038
Qy
4749 TGAGAAATCTGAGCGCTATTTTATTTATCAGTGAAGTCACTCAGGAGATCTCTCTAC 4808
Db
5039 TGAGAAATCTGAGCGCTATTTTATTTATCAGTGAAGTCACTCAGGAGATCTCTCTAC 5098
Qy
4809 GCCGACGATCTGTGGCGGACCTGCAGGTCCGCATCAACCGCGCAACAGTGTGGTGTCT 4868
Db
5099 GCCGACGATCTGTGGCGGACCTGCAGGTCCGCATCAACCGCGCAACAGTGTGGTGTCT 5158
Qy
4869 GCGGCTATATCGCGACATCAACGATGGGAAGATCGGCTCGCACTTCGGGCTCATG 4928
Db
5159 GCGGCTATATCGCGACATCAACGATGGGAAGATCGGCTCGCACTTCGGGCTCATG 5218
Qy
4929 AGCGCTTGTTCGGGCTGGGTATGTGTGCGAGCCCGCTGGCGGGGAGACTGTGGGGCC 4988
Db
5219 AGCGCTTGTTCGGGCTGGGTATGTGTGCGAGCCCGCTGGCGGGGAGACTGTGGGGCC 5278
Qy
4989 ATCTCTCTTG----- 4997
Db
5279 ATCTCTCTTGACCTGCAGGGGGGGGGGAAAGCCACGTTGTGTCTCAAAATCTCTGAT 5338
Qy
4998 ----- 4997
Db
5339 GTTACATTGCACAGATAAATAATATATCATATGACAAATAAATCTGCTTACATAA 5398
Qy
4998 ----- 4997
Db
5399 ACAGTAATACAGGGGTGTATGAGCCATATTCACGGGAAACGTCCTTCTCAAGGGCCG 5458
Qy
4998 ----- 4997
Db
5459 GATTAAATTCACATGATGCTGATTTATATGGGTATTAATGGGCTCGGATTAATGTG 5518
Qy
4998 ----- 4997
Db
5519 GGCAATCAGGTGCGACAACTTATCGATTGTATGGAAGCCGATGCGCGAGTTGTTTC 5578
Qy
4998 ----- 4997

5579 TGAACATGCGAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACT 5638
Qy
4998 ----- 4997
Db
5639 GGCTGACGGAATTTATGCTCTCTCCGACCATCAAGCATTTTATCCGTACTCTCTGATGATG 5698
Qy
4998 ----- 4997
Db
5699 CATGTTACTCACCACCTGCGATCCCGGGAACACAGCATTCACAGGTATTAGAAGATATC 5758
Qy
4998 ----- 4997
Db
5759 CTGATTGAGTGAAAAATATTGTTGATCGCTGGCAGTCTCTCGCGGGTTCATTCGA 5818
Qy
4998 ----- 4997
Db
5819 TTCTCTGTTTGAATTTCTCTTTTAAACAGCGATCGGTATTTCTCTCGCTCAGCGCAAT 5878
Qy
4998 ----- 4997
Db
5879 CACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGCTGGC 5938
Qy
4998 ----- 4997
Db
5939 CTGTTGAACAGTCTGGAAGAAATGCATAAGCTTTGCCATTTCTCACCGGATTCAGTCG 5998
Qy
4998 ----- 4997
Db
5999 TCACCTCATGGTGAATTTCTCTCTCAATTAACAGAACTTATTTTGACGAGGGAAATTAATAGTT 6058
Qy
4998 ----- 4997
Db
6059 GTATTGATGTTGACAGTTCGGAATCGAGACCGATACAGGATCTTGCCATCTCTATGGA 6118
Qy
4998 ----- 4997
Db
6119 ACTGCTCGGTGAGTTTCTCTCTCAATTAACAGAAACGCTTTTTCAAAAATATGATTTG 6178
Qy
4998 ----- 4997
Db
6179 ATAATCTGATATGAATAAATGTCAGTTTCAATGATGCTCGATGAGTTTCTTAATCAG 6238
Qy
4998 ----- 4997
Db
6239 AATGTTAATTTGGTTGTAACATGCGAGAGCATTTACGCTTGACGGGACGCGGCT 6298
Qy
4998 ----- 4997
Db
6299 TTGTTCAATAAATCGAACTTTTGTGCTGAGTTGAAGATCAGATCAGCATCTTCCGACAA 6358
Qy
4998 ----- 4997
Db
6359 CGCAGACCGTTCCGTGGCAAGAAAAAGTTCAAAATCAACAACTGCTCCACCTACAACAA 6418
Qy
4998 ----- 4997
Db
6419 AGCTCTCATNAACCGTGGCTCCCTCACTTTCTGGCTGGATGATGGGGGATTCAGGCCGTG 6478
Qy
4998 ----- 4997
Db
6479 GTATGATCAGCAACACCTTTCTTTCAGAGGACAGACTCAGCGCCCTCCCTCTGAGG 6538
Qy
5009 CCTTGGCGGGCGGTCTCAACGGCTCAACCTACTACTGGGCTGCTTCTCTTAATCAGGA 5068
Db
6539 TCCACGCGGGGGGTCTCAACGGCTCAACCTACTACTGGGCTGCTTCTCTTAATCAGGA 6598
Qy
5069 GTCCGATAAGGAGAGCGTCTGAGTATCTATGATGGAAGTATGGGAATGGTATACCCGC 5128
Db
6599 GTCCGATAAGGAGAGCGTCTGAGTATCTATGATGGAAGTATGGGAATGGTATACCCGC 6658
Qy
5129 ATTCTTCAGTGTGAGGTCTCTCTATCAGATTTGCTCCCACTTAAGCAACCGGAGG 5188
Qy

Db 6659 ATTCTTCAGTCTCTGAGGCTCTCTATCAGANTANGCCCAACTAAAGCAACCGGAGAGG 6718
Qy 5189 AGATTTTCATGTAATTTCTCTGACTTTTGGTCTATCAGTAGACTCGAAGTGTGAGACTAT 5248
Db 6719 AGATTTTCATGTAATTTCTCTGACTTTTGGTCTATCAGTAGACTCGAAGTGTGAGACTAT 6778
Qy 5249 CTCGGTTATGACAGGAGAAATGTCCTTCTTGAGACAGTAATGAAGTCCCAACCAATAAA 5308
Db 6779 CTCGGTTATGACAGGAGAAATGTCCTTCTTGAGACAGTAATGAAGTCCCAACCAATAAA 6838
Qy 5309 GAAATCCTTTGTTATCAGGAAACAAACTCTCTGTTTGGAACTTTTTCGGTGCCTTGAACATAT 5368
Db 6839 GAAATCCTTTGTTATCAGGAAACAAACTCTCTGTTTGGAACTTTTTCGGTGCCTTGAACATAT 6898
Qy 5369 AAAATGTAGAGTGATATGTCGGGTAGGAAATGAGAGCGGCAAAATGCTTACCTTCGGAAC 5428
Db 6899 AAAATGTAGAGTGATATGTCGGGTAGGAAATGAGAGCGGCAAAATGCTTACCTTCGGAAC 6958
Qy 5429 TTCAGAGGTATGAGGTTTGTAGATACATGCAACTTTCAGTGACCAAGTGTCTATT 5488
Db 6959 TTCAGAGGTATGAGGTTTGTAGATACATGCAACTTTCAGTGACCAAGTGTCTATT 7018
Qy 5489 TCGTTCAAACCAATCCGAATCCAGAGAAATCAAAGTTTGTCTACTATTGATCCAAAGC 5548
Db 7019 TCGTTCAAACCAATCCGAATCCAGAGAAATCAAAGTTTGTCTACTATTGATCCAAAGC 7078
Qy 5549 CAGTCGGTCTTGAACCTGCAATAGTGTCTGCTGTTTTCAGGTCATCTTTGATGAAT 5608
Db 7079 CAGTCGGTCTTGAACCTGCAATAGTGTCTGCTGTTTTCAGGTCATCTTTGATGAAT 7138
Qy 5609 AAATCTAGTCTTTGATCTAAATTAATTTGAGCGCAAGGCGATAAATACCAAACTAA 5668
Db 7139 AAATCTAGTCTTTGATCTAAATTAATTTGAGCGCAAGGCGATAAATACCAAACTAA 7198
Qy 5669 AACTCTTTTAAACGTTAAAGGCAAGTATGTCCTGTATTAATTAACCCCAATTCAGCT 5728
Db 7199 AACTCTTTTAAACGTTAAAGGCAAGTATGTCCTGTATTAATTAACCCCAATTCAGCT 7258
Qy 5729 CGTAGCTGATCTCAATCAACTGAGGCGCACTATCTTTGTTTGAAGAAATTCGGGAGA 5788
Db 7259 CGTAGCTGATCTCAATCAACTGAGGCGCACTATCTTTGTTTGAAGAAATTCGGGAGA 7318
Qy 5789 TCGCATATCGAGAAAGGTCAGCTGATTTTAAACGTAAGTAAATTCATCAAGATCTCTGC 5848
Db 7319 TCGCATATCGAGAAAGGTCAGCTGATTTTAAACGTAAGTAAATTCATCAAGATCTCTGC 7378
Qy 5849 CTCGCGGTTTCGGTGATGACGGTGAAACCTCTGACACATGCAGCTCCCGGAGCGGTC 5908
Db 7379 CTCGCGGTTTCGGTGATGACGGTGAAACCTCTGACACATGCAGCTCCCGGAGCGGTC 7438
Qy 5909 ACAGCTTGTCTGTAAGCGGATCCGGGAGCAGACAAAGCCGTCAGGCGCGCTCAGCGGGT 5968
Db 7439 ACAGCTTGTCTGTAAGCGGATCCGGGAGCAGACAAAGCCGTCAGGCGCGCTCAGCGGGT 7498
Qy 5969 GTTGGCGGTGTCGGGCGCAGCCATAGCCAGTCAAGTACGAGGATGAGGAGTGTATCT 6028
Db 7499 GTTGGCGGTGTCGGGCGCAGCCATAGCCAGTCAAGTACGAGGATGAGGAGTGTATCT 7558
Qy 6029 GCTTAACTATCGGCATCAGAGCAGATTTGATCAGAGTGCAACCATATCGGCTGTGAA 6088
Db 7559 GCTTAACTATCGGCATCAGAGCAGATTTGATCAGAGTGCAACCATATCGGCTGTGAA 7618
Qy 6089 TACCGCACAGATCGGTAAAGAGAAATACCGCATCAGCGGCTCTTCCGTTCTCTGCTCA 6148
Db 7619 TACCGCACAGATCGGTAAAGAGAAATACCGCATCAGCGGCTCTTCCGTTCTCTGCTCA 7678
Qy 6149 CTGACTCTGCTGCGCTCGGTCTGCTGCGGAGCGGTATCAGCTCACTCAAAAGCGG 6208
Db 7679 CTGACTCTGCTGCGCTCGGTCTGCTGCGGAGCGGTATCAGCTCACTCAAAAGCGG 7738
Qy 6209 TAATACGGTTATCCACAGAAATCAGGGGATAAGCGCAGGAAAGCAATGTGAGCAAAAGGCC 6268
Db 7739 TAATACGGTTATCCACAGAAATCAGGGGATAAGCGCAGGAAAGCAATGTGAGCAAAAGGCC 7798

Qy 6269 AGCAAAAGGCCAGGAAACCGTAAAGAGCGCGTTCCTGCTGGCGTTTTCATAGGCTCCGCC 6328
Db 7799 AGCAAAAGGCCAGGAAACCGTAAAGAGCGCGTTCCTGCTGGCGTTTTCATAGGCTCCGCC 7858
Qy 6329 CCCTCAGCAGCATCAAAATTCAGAGCTCAAGTCAGAGGTGGGAAACCCGACAGGAC 6388
Db 7859 CCCTCAGCAGCATCAAAATTCAGAGCTCAAGTCAGAGGTGGGAAACCCGACAGGAC 7918
Qy 6389 TATAAAGATACAGAGGCTTCCCTCGGAGCTCCCTGAGCTCTGTCGCTCTCTGTTCCGACCC 6448
Db 7919 TATAAAGATACAGAGGCTTCCCTCGGAGCTCCCTGAGCTCTCTGTTCCGACCC 7978
Qy 6449 TCGCGCTTACCGGATACCTGCGCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCAT 6508
Db 7979 TCGCGCTTACCGGATACCTGCGCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCAT 8038
Qy 6509 GCTCAAGCTGTAGGTATCTCAGTTCCGCTGAGTGTCTGCTCCAGCTGGGCTGTGTC 6568
Db 8039 GCTCAAGCTGTAGGTATCTCAGTTCCGCTGAGTGTCTGCTCCAGCTGGGCTGTGTC 8098
Qy 6569 ACGAACCCCGCTTACGCGGAGCGCTGCGCTTATCCGCTTAACTATCTCTTGGTCCA 6628
Db 8099 ACGAACCCCGCTTACGCGGAGCGCTGCGCTTATCCGCTTAACTATCTCTTGGTCCA 8158
Qy 6629 ACCCGTAAGACACAGCTTATCGCCACTGCGCAGCAGCCACTGCTAACAGGATAGCAGAG 6688
Db 8159 ACCCGTAAGACACAGCTTATCGCCACTGCGCAGCAGCCACTGCTAACAGGATAGCAGAG 8218
Qy 6689 CGAGGTATGTAGGCGGTGTACAGATCTTGAAGTGGTGGCTTAACTACAGGCTACACTA 6748
Db 8219 CGAGGTATGTAGGCGGTGTACAGATCTTGAAGTGGTGGCTTAACTACAGGCTACACTA 8278
Qy 6749 GAAGGACAGTATTTGGTATCTGCGCTCTCTGAAGCCAGTTACCTTCGGAAGAAAGATTG 6808
Db 8279 GAAGGACAGTATTTGGTATCTGCGCTCTCTGAAGCCAGTTACCTTCGGAAGAAAGATTG 8338
Qy 6809 GTAGCTCTTGAATCCGCAAAACCAACCAACCGCTGGTAGCGGTGTTTTTTTGTGCAAGC 6868
Db 8339 GTAGCTCTTGAATCCGCAAAACCAACCAACCGCTGGTAGCGGTGTTTTTTTGTGCAAGC 8398
Qy 6869 AGCAGATTACGCGCAGAAAGGATCTCAGAGATCCTTTTGAATCTTTTCTACCGGGT 6928
Db 8399 AGCAGATTACGCGCAGAAAGGATCTCAGAGATCCTTTTGAATCTTTTCTACCGGGT 8458
Qy 6929 CTGACGCTCAGTGAAGCAAGAACTCAGCTTAAAGGATTTTGGTTCATGAGATTATCAAAA 6988
Db 8459 CTGACGCTCAGTGAAGCAAGAACTCAGCTTAAAGGATTTTGGTTCATGAGATTATCAAAA 8518
Qy 6989 GATCTTTCAGCTAGATCTCTTTTAAATTAAGATTTTAAATCAATCTAAAGTATAT 7048
Db 8519 GATCTTTCAGCTAGATCTCTTTTAAATTAAGATTTTAAATCAATCTAAAGTATAT 8578
Qy 7049 ATGAGTAAACTTCTGTCAGCTTACCAATGCTTAAATCAGTGAGGCACTTATCTCAGCGA 7108
Db 8579 ATGAGTAAACTTCTGTCAGCTTACCAATGCTTAAATCAGTGAGGCACTTATCTCAGCGA 8638
Qy 7109 TCTGCTTATTTCTGTCATCATAGTGTGCTGATCTCCCGCTGCTGATAGTAACCTACGATAC 7168
Db 8639 TCTGCTTATTTCTGTCATCATAGTGTGCTGATCTCCCGCTGCTGATAGTAACCTACGATAC 8698
Qy 7169 GGGAGGCTTACCATCTGCGCCAGTGTGCAATGATACCGGAGACCCACGCTCACCGG 7228
Db 8699 GGGAGGCTTACCATCTGCGCCAGTGTGCAATGATACCGGAGACCCACGCTCACCGG 8758
Qy 7229 CTCAGATTTTATCAGCAATTAACCCAGCAGCCGGAAGGCGGAGCGAGAGTGGTCTG 7288
Db 8759 CTCAGATTTTATCAGCAATTAACCCAGCAGCCGGAAGGCGGAGCGAGAGTGGTCTG 8818
Qy 7289 CAATCTTATCGGCTCCATCAGTCTATTAATTTGTCGCGGAGCTAGAGTAGTAGTT 7348
Db 8819 CAATCTTATCGGCTCCATCAGTCTATTAATTTGTCGCGGAGCTAGAGTAGTAGTT 8878

7349 GCCAGTTAAAGTTGGCAACGTTGTTGTCATTGTCGAGGCATCGTGTGTCAGCT 7408
Db 8879 GCCAGTTAAAGTTGGCAACGTTGTTGTCATTGTCGAGGCATCGTGTGTCAGCT 8938
Qy 7409 CGTCTTTGGTATGGCTTCATCAGTCCGCTCCCAACGATCAAGCGGAGTTACATGAT 7468
Db 8939 CGTCTTTGGTATGGCTTCATCAGTCCGCTCCCAACGATCAAGCGGAGTTACATGAT 8998
Qy 7469 CCCCATCTGTCGAAACGAGTGTAGCTCTCTTCGCTGCTCCGATCGTTGTCAGAGTA 7528
Db 8999 CCCCATCTGTCGAAACGAGTGTAGCTCTCTTCGCTGCTCCGATCGTTGTCAGAGTA 9058
Qy 7529 AGTTGGCGCAGTGTATCATCTCATGTTATGGCAGCAGTCAATATCTCTTACTGTCA 7588
Db 9059 AGTTGGCGCAGTGTATCATCTCATGTTATGGCAGCAGTCAATATCTCTTACTGTCA 9118
Qy 7589 TGCCATCCGTAAGATGCTTTCTCTGAGTGTGAGTACTCAACCAAGTCAATCTGAGAAT 7648
Db 9119 TGCCATCCGTAAGATGCTTTCTCTGAGTGTGAGTACTCAACCAAGTCAATCTGAGAAT 9178
Qy 7649 AGTGTATCGCGCGACCGAGTTGCTCTTTCGCGCGGTCAACACGGGATAATACCGCGCCAC 7708
Db 9179 AGTGTATCGCGCGACCGAGTTGCTCTTTCGCGCGGTCAACACGGGATAATACCGCGCCAC 9238
Qy 7709 ATAGCAGAACTTTAAAGTGTCTCATCTGGAACAGTCTTCGCGCGGAAACTCTCAA 7768
Db 9239 ATAGCAGAACTTTAAAGTGTCTCATCTGGAACAGTCTTCGCGCGGAAACTCTCAA 9298
Qy 7769 GATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCACTCGTCGACCAACTGATCTT 7828
Db 9299 GATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCACTCGTCGACCAACTGATCTT 9358
Qy 7829 CAGCATCTTTTACTTTCACGAGCTTTCTGGGTGAGCAAAACAGGAAGGCAAAATCCG 7888
Db 9359 CAGCATCTTTTACTTTCACGAGCTTTCTGGGTGAGCAAAACAGGAAGGCAAAATCCG 9418
Qy 7889 CAAAAAGGGAATAGGCGCAGACGGAATGTCATCACTACTACTCTCTCTTTTCAAT 7948
Db 9419 CAAAAAGGGAATAGGCGCAGACGGAATGTCATCACTACTACTCTCTCTTTTCAAT 9478
Qy 7949 ATTATTGAAGCATTTATCAGGGTTATTGTTCTCATGCGGATACATATTTCAATGATTT 8008
Db 9479 ATTATTGAAGCATTTATCAGGGTTATTGTTCTCATGCGGATACATATTTCAATGATTT 9538
Qy 8009 AGAAAAATAAACAATAGGGTTCGCGCAGATTTCCCGAAGTCCGACCTGACGTCT 8068
Db 9539 AGAAAAATAAACAATAGGGTTCGCGCAGATTTCCCGAAGTCCGACCTGACGTCT 9598
Qy 8069 AAGAAACCATTTATCATGATTAACCTATAAATAAGCGGTATCAGGAGCCCTTTC 8128
Db 9599 AAGAAACCATTTATCATGATTAACCTATAAATAAGCGGTATCAGGAGCCCTTTC 9658
Qy 8129 GTCTTCAAGATTAATTCCTCATGTTGACGCTTATCATGATGAGTCAATGTTG 8188
Db 9659 GTCTTCAAGATTAATTCCTCATGTTGACGCTTATCATGATGAGTCAATGTTG 9718
Qy 8189 TATTGTGAATAGACGAGATCGGACACTGGAATAAATACAGTTATTATTCG 8241
Db 9719 TATTGTGAATAGACGAGATCGGACACTGGAATAAATACAGTTATTATTCG 9771

RESULT 11
LOCUS AY178634 9139 bp DNA circular SYN 06-APR-2003
DEFINITION Expression vector pPYX99, complete sequence.
ACCESSION AY178634
VERSION AY178634.1 GI:29568880
KEYWORDS Expression vector pPYX99
ORGANISM Expression vector pPYX99
SOURCE artificial sequences; vectors.
REFERENCE 1 (bases 1 to 9139)
AUTHORS Yao, Q., Peng, R. and Xiong, A.

TITLE Direct Submission
JOURNAL Submitted (12-NOV-2002) Shanghai Academy of Agricultural Sciences,
Agro-Biotechnology Research Center, 2901 Beidi Rd, Shanghai 201106,
China
COMMENT NCBI staff are still waiting for submitters to provide appropriate
coding region information.
FEATURES
Location/Qualifiers
source
1..9139
/organism="Expression vector pPYX99"
/mol_type="genomic DNA"
/db_xref="taxon:218468"
/lab_host="Pichia pastoris"
misc_feature
1..936
/note="modified AOX1 promoter"
misc_feature
937..1212
/note="similar to modified alpha factor secretion signal"
terminator
1275..1607
/note="AOX1 transcription terminator"
misc_feature
complement(1994..4527)
/note="similar to phosphoribosyl-ATP pyrophosphohydrolase"
CDS
4798..5613
/codon_start=1
/product="aminoglycoside-3'-O-phosphotransferase"
/protein_id="AA084033.1"
/db_xref="GI:29568881"
/translation="MSHIQRETSRPLNSNDADLYGKWARDNVQSGATVRLY
GKPAPELFLKHGKSVANDVTDEMRLNWLTEFMPFTIKHFIPTDDAWLLTALP
KTFAPLEELYDPSGSENIYDALAVFLRLHLSIPVNCPCFNSDRVFLQAQSRNNGL
VDASFDDERNGWPEQVQKEMHKLFPSPDSVTHGDFSLDNLIFDEGLIGCIDVG
RVGIADRYQDLAILWNCLEGFPSLQKLFQKYGIDNPDMNKLOFHLMDLDF"
5987..6736
misc_feature
/note="3' AOX1 fragment from Pichia pastoris"
rep_origin
complement(7151..7824)
/note="ColE1"
CDS
complement(7969..8829)
/function="ampicillin resistance"
/codon_start=1
/product="beta-lactamase"
/protein_id="AA084034.1"
/db_xref="GI:29568882"
/translation="MSIOHFRVALIPFAARCLPVFAHPETLVKVKDAEDOLGARVGY
IELDLSKILLESFPEPFPMSTFVKVLCGAVLSVDAGEQGLRRHIVSQNDLVE
YSPVTERKLTDMGTVELCSAAITMSDNTAANLLLTIGGPKELTFLHNGDVRRL
DRWEPELNEAIPNDERDTPAAMATTLIRKLLTGLLTLASRQQLIDWMEADKVGAPL
LRSLALPWGFIADKSGAGERSGLIIAALGPDGKPSRIVTVITGTSQATMDERNQIA
EIGASLIKHW"

ORIGIN

Query Match 75.8%; Score 6250.2; DB 12; Length 9139;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 7817; Conservative 0; Mismatches 228; Indels 1290; Gaps 14;
Qy 1 AGATCTAATCTCAAGCAGAAAGGTTGAATCAACACCTTTTGGCATCCGACATCCACAG 60
Db 1 AGATCTAATCTCAAGCAGAAAGGTTGAATCAACACCTTTTGGCATCCGACATCCACAG 60
Qy 61 GTCCATTCTCACATAAGTGCACAAACCAACAGAGGGGATACACTAGCAGAGACCGT 120
Db 61 GTCCATTCTCACATAAGTGCACAAACCAACAGAGGGGATACACTAGCAGAGACCGT 120
Qy 121 TCGAAACGAGGACCTCCACTCTCTTCTCTACACCCACTTTTGGCATCGAAACACC 180
Db 121 TCGAAACGAGGACCTCCACTCTCTTCTCTACACCCACTTTTGGCATCGAAACACC 180
Qy 181 AGCCAGATTATTGGGCTTGAATGGAGCTCGCTCAATTCCTTCTATTAGGCTACTA 240
Db 181 AGCCAGATTATTGGGCTTGAATGGAGCTCGCTCAATTCCTTCTATTAGGCTACTA 236
Qy 241 ACACCATGACTTTATAGCTGTCTATCTGCGCCCTCGGAGGTTCAATGTTGTTA 300
Db 237 ACACCATGACTTTATAGCTGTCTATCTGCGCCCTCGGAGGTTCAATGTTGTTA 296
Qy 301 TTTCCGATGCAACAGCTCCGCATTACACCCGACATCACTCCAGATGAGGCTTCTG 360

Db 297 TTTCCGAATGCAACAAAGCTCGCATTAACACCGAATCACTCCAGATGAGGCTTTCTG 356
Qy 361 AGTGTGGGTCAAAATAGTTTCATGTTCCCAAAATGGCCCAAACTGACAGTTTAAACGCT 420
Db 357 AGTGTGGGTCAAAATAGTTTCATGTTCCCAAAATGGCCCAAACTGACAGTTTAAACGCT 416
Qy 421 GTCTTGGAACTTAATATGACAAAAGCGTGATCTCATCCAGATGAACCTAAGTTTGGTTG 480
Db 417 GTCTTGGAACTTAATATGACAAAAGCGTGATCTCATCCAGATGAACCTAAGTTTGGTTG 476
Qy 481 TTGAAATGCTAAACCGGTCAGTTTGGTCAAAAAGAACTTCCAAAAGTCGCCATACCGTTTGT 540
Db 477 TTGAAAATGCTAACCGGTCAGTTTGGTCAAAAAGAACTTCCAAAAGTCGCCATACCGTTTGT 536
Qy 541 CTGTGTTGGTATGATGACAAATGCTCAAAAATTAATCTCATTAATGCTTAGCGCAGTCT 600
Db 537 CTGTGTTGGTATGATGACAAATGCTCAAAAATTAATCTCATTAATGCTTAGCGCAGTCT 596
Qy 601 CTCTATCGCTTCTGAAACCGGTCACCTGTGCGGAACGCAATGGGAAACACCGCT 660
Db 597 CTCTATCGCTTCTGAAACCGGTCACCTGTGCGGAACGCAATGGGAAACACCGCT 656
Qy 661 TTTTGGATGATTAATGCAATGCTCCACATTTGTATGCTTCCAAGATTCTGGTGGGAATACT 720
Db 657 TTTTGGATGATTAATGCAATGCTCCACATTTGTATGCTTCCAAGATTCTGGTGGGAATACT 716
Qy 721 GCTGATAGCTAAAGTTCATGATCAAAATTTAACTGTCTAACCCCTACTTGACAGCAAT 780
Db 717 GCTGATAGCTAAAGTTCATGATCAAAATTTAACTGTCTAACCCCTACTTGACAGCAAT 776
Qy 781 ATATAACAGAGGAAGTGCCTGTCTTAAACCTTTTATCATCATTAATAGCTT 840
Db 777 ATATAACAGAGGAAGTGCCTGTCTTAAACCTTTTATCATCATTAATAGCTT 836
Qy 841 ACTTTCATPAATGGCACTGGTTCATTAATGACAAGCTTTGATTTAAGCACTTTAAACGA 900
Db 837 ACTTTCATPAATGGCACTGGTTCATTAATGACAAGCTTTGATTTAAGCACTTTAAACGA 896
Qy 901 CAACCTTGAGAGATCAAAAACACTAATTTTGAAGGATCCAAAGATGAGATTTCCT 960
Db 897 CAACCTTGAGAGATCAAAAACACTAATTTTGAAGGATCCAAAGATGAGATTTCCT 951
Qy 961 TCAATTTTATGTCAGTTTATTCGACATCTCCGCAATAGCTGTCTCAGTCAACACT 1020
Db 952 TCTATCTTCACTCGCTGTCTTGTTCGCTCTTCTGCTTGGCTGCTCCTGTCAACACT 1011
Qy 1021 ACAACAGAGATGAACCGCAAAATTCGGCTGAGCTGTCTATCGGTTACTCAGATTTA 1080
Db 1012 ACTACCGAGGATGAACCTGCTCAAAATCCTGCTGAGGCTGTCTGCTGCTGCTGCTG 1071
Qy 1081 GAAGGGGATTCGATGTTGCTGTTTGGCATTTTCCAAACAGACAAAATAACGGGTTATTG 1140
Db 1072 GAGGGTGACTTCGAGCTGCTGCTTGGCATTTCTCTAACTCCACCAACACGGTTTGTG 1131
Qy 1141 TTTATAATACTACTATTGCCAGCATTTGCTGCTAAAGAGAGGGGTATCTCTCGAGAAA 1200
Db 1132 -----GAGGAGGCTGAAGCTGAAGCTGAAGCTGAAGCTTAAATTTTCAACACT 1173
Qy 1201 AGAGGGCTGAAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 1260
Db 1174 ACTATGCTTCTATGCTGCTTAAGGAGGAGGTTTCCCTCGAGAAAAGAGAGGCTGAG 1233
Qy 1261 GCGCCCGGGGAGGAGGATAAATTTGGCTTTCCCGGTGTCAACCGCCAGCAGTGCACGGAG 1320
Db 1234 GCTTACGTAGATTGA----- 1249
Qy 1321 AGAGGTTGCTGTTTGTATGACAGTGTCCGGGATTTCCCGGTGTCTCCACCCCATGGCC 1380
Db 1250 ----- 1249
Qy 1381 ATCGAGAACACTCAAGAGAGAGATGTCCTCTTAACTAGTAGGCGTAGAATTTCCCTAGGG 1440

Db 1250 -----GCTCGGG 1256
Qy 1441 CCGCGCGAATTAATTCGCTTAGACATGACTGTTCTCCTCAGTTCAAGTTGGGCATTTAGC 1500
Db 1257 CCGCGCGAATTAATTCGCTTAGACATGACTGTTCTCCTCAGTTCAAGTTGGGCATTTAGC 1316
Qy 1501 AGAAGACCGGCTCTGCTAGATTCTAATCAAGAGGATGTCAAGATGCCATTTCCCTGAGAG 1560
Db 1317 AGAAGACCGGCTCTGCTAGATTCTAATCAAGAGGATGTCAAGATGCCATTTCCCTGAGAG 1376
Qy 1561 ATGCAAGGCTCAATTTTGTAGTCTTTTATTTGTAACTATATAGTATAGGATTTTTTTT 1620
Db 1377 ATGCAAGGCTCAATTTTGTAGTCTTTTATTTTGTAACTATATAGTATAGGATTTTTTTT 1436
Qy 1621 GTCATTTGCTTCTCTGCTAGAGCTTCTCTCATGATCAGCCTATCTCGACGCTCATGAA 1680
Db 1437 GTCATTTGCTTCTCTGCTAGAGCTTCTCTCATGATCAGCCTATCTCGACGCTCATGAA 1496
Qy 1681 TATCTTGTGGTAGGGTTTGGGAAATCATTCGAGTTTGATGTTTTTCTTTGGTATTTTCC 1740
Db 1497 TATCTTGTGGTAGGGTTTGGGAAATCATTCGAGTTTGATGTTTTTCTTTGGTATTTTCC 1556
Qy 1741 ACTCCTCTTCAGAGTACAGAAATTAAGTGAGAGTTCTGTTTGTGCAAGCTTATCGATAA 1800
Db 1557 ACTCCTCTTCAGAGTACAGAAATTAAGTGAGAGTTCTGTTTGTGCAAGCT-----A 1608
Qy 1801 GCTTTAATCGGCTAGTTTATCAAGTTAAATTTGCTAAACGAGTCAGGCCACCGTATGAA 1860
Db 1609 GCTTTAATCGGCTAGTTTATCAAGTTAAATTTGCTAAACGAGTCAGGCCACCGTATGAA 1668
Qy 1861 ATCTAAATGGGCTCATCTCGACATCTCGGACCGTCAACCTGGATGCTGTAGGCAATAG 1920
Db 1669 ATCTAAATGGGCTCATCTCGACATCTCGGACCGTCAACCTGGATGCTGTAGGCAATAG 1728
Qy 1921 CTTGGTTATGCGGTAAGTCTCGGCGCTCTTGGGGATATCGTCCATTCGACAGCATCGC 1980
Db 1729 CTTGGTTATGCGGTAAGTCTCGGCGCTCTTGGGGATATCGTCCATTCGACAGCATCGC 1788
Qy 1981 CAGTCACTATGCGCTCTCTAGCGCTATATGCTTGTGATGCAATTTCTATAGGCAACCGT 2040
Db 1789 CAGTCACTATGCGCTCTCTAGCGCTATATGCTTGTGATGCAATTTCTATAGGCAACCGT 1848
Qy 2041 TCTCGGAGCACTGTGCGACCGCTTTGGCGCGCCGCGCAGTCTGCTCGCTTGGTACTTGG 2100
Db 1849 TCTCGGAGCACTGTGCGACCGCTTTGGCGCGCCGCGCAGTCTGCTCGCTTGGTACTTGG 1908
Qy 2101 AGCCACTATGCACTAGCGGATCATGCGACCAACCCGCTCTGTGGATCTATCGAATCTA 2160
Db 1909 AGCCACTATGCACTAGCGGATCATGCGACCAACCCGCTCTGTGGATCTATCGAATCTA 1968
Qy 2161 AATGTAAGTTAAATCTTAATAATTAATTAAGTCCGAGTTTCTCCATACGAACTTAA 2220
Db 1969 AATGTAAGTTAAATCTTAATAATTAATTAAGTCCGAGTTTCTCCATACGAACTTAA 2028
Qy 2221 CAGCATTTGCGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAA 2280
Db 2029 CAGCATTTGCGGTGAGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAA 2088
Qy 2281 TATGTTTCACTCCCTCAGGAGTTACGTTTGTGAAGTGAATCTTCTGGAAGGTTGCGAG 2340
Db 2089 TATGTTTCACTCCCTCAGGAGTTACGTTTGTGAAGTGAATCTTCTGGAAGGTTGCGAG 2148
Qy 2341 TGTTAATCCGCTGATTTGACGGGATATCCGTAAGTGTGCAAGGTTGTTGTTGTTACCGG 2400
Db 2149 TGTTAATCCGCTGATTTGACGGGATATCCGTAAGTGTGCAAGGTTGTTGTTGTTACCGG 2208
Qy 2401 AGGAGTAATCTCCAACTCTCTGGAGAGTAGGCAACAAACAAACACAGATCCAGCGTGT 2460
Db 2209 AGGAGTAATCTCCAACTCTCTGGAGAGTAGGCAACAAACAAACACAGATCCAGCGTGT 2268
Qy 2461 GTACTTGAATCAATAAGAGAGCAATTTCTCGATTTGCGAGTCAAGTGTTCAGGAGCGT 2520
Db 2269 GTACTTGAATCAATAAGAGAGCAATTTCTCGATTTGCGAGTCAAGTGTTCAGGAGCGT 2328

QY 2521 ACTGATTCGACATTTCCAAAGCCTGCTCGTAGGTTGCAACCGATAGAGGTTGTAGAGTGTG 2580
DB |||||
QY 2329 ACTGATTCGACATTTCCAAAGCCTGCTCGTAGGTTGCAACCGATAGAGGTTGTAGAGTGTG 2388
DB |||||
QY 2581 CAATACACCTTGGGTACAAATTTCAACCCCTTGGCAACTGCACAGCTTGGTTGTGAACAGCAT 2640
DB |||||
QY 2389 CAATACACCTTGGGTACAAATTTCAACCCCTTGGCAACTGCACAGCTTGGTTGTGAACAGCAT 2448
DB |||||
QY 2641 CTTCAATTTGGCAAGCTTCTGTCTGTATATCGACAGCCAAACAGAAATCACTTGGGAAT 2700
DB |||||
QY 2449 CTTCAATTTGGCAAGCTTCTGTCTGTATATCGACAGCCAAACAGAAATCACTTGGGAAT 2508
DB |||||
QY 2701 CAATACCACTTGCAGCTTGAGACAGAGGCTCTGAGGCAACGAAATCTGGATCAGCGTATT 2760
DB |||||
QY 2509 CAATACCACTTGCAGCTTGAG - CAGAAGCTCTGAGGCAACGAAATCTGGATCAGCGTATT 2567
DB |||||
QY 2761 TATCAGCAATTAACAGAACTTGCAGAGGCCCGACAGGCGATGTCAATACTACAGAGGCTG 2820
DB |||||
QY 2568 TATCAGCAATTAACAGAACTTGCAGAGGCCCGACAGGCGATGTCAATACTACAGAGGCTG 2627
DB |||||
QY 2821 ATGTCTCATTTTGAACCATCATCTTGGCAGCAGTAACGAACTGTCTTCCCTGGACCAATA 2880
DB |||||
QY 2628 ATGTCTCATTTTGAACCATCATCTTGGCAGCAGTAACGAACTGTCTTCCCTGGACCAATA 2887
DB |||||
QY 2881 TTTTGTCACTTAGGAACAGTTTCTGTTCGTAAAGCCATAGCAGCTACTGCCTGGCGCC 2940
DB |||||
QY 2698 TTTTGTCACTTAGGAACAGTTTCTGTTCGTAAAGCCATAGCAGCTACTGCCTGGCGCC 2747
DB |||||
QY 2941 CTCCTGCTAGCAGNATACACTTAGCACCACCACTTGTGGGCAACGATAGACTTCTGGGG 3000
DB |||||
QY 2748 CTCCTGCTAGCAGNATACACTTAGCACCACCACTTGTGGGCAACGATAGACTTCTGGGG 2807
DB |||||
QY 3001 TAAAGGTAACCTTCTTCTTAGGTGAGATGCAAAAACAAATTTCTTTCGAACCCAGCAACTT 3060
DB |||||
QY 2808 TAAAGGTAACCTTCTTCTTAGGTGAGATGCAAAAACAAATTTCTTTCGAACCCAGCAACTT 2867
DB |||||
QY 3061 TGGCAGGAACCCAGCATCAGGAGTGGAGGAGAGATTTGGCGTTCCACAGGAATAT 3120
DB |||||
QY 2868 TGGCAGGAACCCAGCATCAGGAGTGGAGGAGAGATTTGGCGTTCCACAGGAATAT 2927
DB |||||
QY 3121 AGAGGCCAACTTCTCAATAGTCTTTCGAAACGAGAGCAGACTACACAGGCGCAAGTCT 3180
DB |||||
QY 2928 AGAGGCCAACTTCTCAATAGTCTTTCGAAACGAGAGCAGACTACACAGGCGCAAGTCT 2987
DB |||||
QY 3181 CAACTTGCACAGCTTCTTCTTAGTTGAGCTTCATGGAAATTTCTCGACGTTATCTATAGAGA 3240
DB |||||
QY 2988 CAACTTGCACAGCTTCTTCTTAGTTGAGCTTCATGGAAATTTCTCGACGTTATCTATAGAGA 3047
DB |||||
QY 3241 GATCAATGGCTCTCTTAACGTTATCTGCAATTCGCAATTTCTCTGGGAAAGGAGCTT 3300
DB |||||
QY 3048 GATCAATGGCTCTCTTAACGTTATCTGCAATTCGCAATTTCTCTGGGAAAGGAGCTT 3107
DB |||||
QY 3301 CTAAACAGGTTCTTCTTAAAGCGACTCCATCAAACTTGGCAGTTAGTTCTTAAAGGCGCTT 3360
DB |||||
QY 3108 CTAAACAGGTTCTTCTTAAAGCGACTCCATCAAACTTGGCAGTTAGTTCTTAAAGGCGCTT 3167
DB |||||
QY 3361 TGTCAACATTTGAGACATTTGCAATTTGCTGCAATTTGCTGCAATTTGCTGCAATTTGCTGCA 3420
DB |||||
QY 3168 TGTCAACATTTGAGCAATTTGCTGCAATTTGCTGCAATTTGCTGCAATTTGCTGCAATTTGCTG 3227
DB |||||
QY 3421 TTTTCTCGATAGGACGCAAGGCGCATCTTCAATTTCTTGTGAGAGGCGCTTGAAGACGT 3480
DB |||||
QY 3228 TTTTCTCGATAGGACGCAAGGCGCATCTTCAATTTCTTGTGAGAGGCGCTTGAAGACGT 3287
DB |||||
QY 3481 CAATTTGCACAAATTCATAGCACTTCAGAGGAGCTTCTTTAGGTTTGGATCTTCTT 3540
DB |||||
QY 3288 CAATTTGCACAAATTCATAGCACTTCAGAGGAGCTTCTTTAGGTTTGGATCTTCTT 3347
DB |||||
QY 3541 TAGGTTGTCTTGGTGATCTTGGCTTGGCATCTCTCTTCTTCTAGTGACCTTTAGGG 3600
DB |||||
QY 3348 TAGGTTGTCTTGGTGATCTTGGCTTGGCATCTCTCTTCTTCTAGTGACCTTTAGGG 3407
DB |||||

QY 3601 ACTTCATATCCAGGTTTCTCTCCACTCGTCCAAAGTCCACCGTACTTGGCAATCTAA 3660
DB |||||
QY 3408 ACTTCATATCCAGGTTTCTCTCCACTCGTCCAAAGTCCACCGTACTTGGCAATCTAA 3467
DB |||||
QY 3661 CTAATGCAAAATAAATAAATAGTCAGACATTTCCAGGCTATATCTTCTTGGATTAGCTT 3720
DB |||||
QY 3468 CTAATGCAAAATAAATAAATAGTCAGACATTTCCAGGCTATATCTTCTTGGATTAGCTT 3527
DB |||||
QY 3721 CTGCAAGTTTCATCAGCTTCTCCCTAAATTTTAGCGTTT - AACAAATCTTCTGCTCAAT 3779
DB |||||
QY 3528 CTGCAAGTTTCATCAGCTTCTCCCTAAATTTTAGCGTTTCAACAAATCTTCTGCTCAAT 3587
DB |||||
QY 3780 AACCGTTTGGTATAAAGAACTTCTGAGCATTTGCTTTCAGATCCCAAAAGTGGCTTCC 3839
DB |||||
QY 3588 AACCGTTTGGTATAAAGAACTTCTGAGCATTTGCTTTCAGATCCCAAAAGTGGCTTCC 3646
DB |||||
QY 3840 ATGGCTCTTAAGACCTTTCATTTGGCCAAACAGGAAGTGGTTCCAAAGTGCAGAAACCA 3899
DB |||||
QY 3647 ATGGCTCTTAAGACCTTTCATTTGGCCAAACAGGAAGTGGTTTCCAAAGTGCAGAAACCA 3706
DB |||||
QY 3900 ACACCTGTTTGTTCACCAACCAAAATTTCAAGCAGTCTCCATCACAATCCAAATCCGATACC 3959
DB |||||
QY 3707 ACACCTGTTTGTTCACCAACCAAAATTTCAAGCAGTCTCCATCACAATCCAAATCCGATACC 3766
DB |||||
QY 3960 AGCAACTTTTGTAGTTGCTCCAGATGTAGCAGCTTTATACCAACACCGTGCAGGAGAT 4019
DB |||||
QY 3767 AGCAACTTTTGTAGTTGCTCCAGATGTAGCAGCTTTATACCAACACCGTGCAGGAGAT 3826
DB |||||
QY 4020 TGGTAGACTCCAGTTTGTGCTTATAGCTTCCGGAATAGACTTTTTCGAGAGGATACACC 4079
DB |||||
QY 3827 TGGTAGACTCCAGTTTGTGCTTATAGCTTCCGGAATAGACTTTTTCGAGAGGATACACC 3886
DB |||||
QY 4080 AGGCCCAACAGTAATAGAGAGTCAGCCACCAAAAGTAGTAATAGACCATTCGGGGGG 4139
DB |||||
QY 3887 AGGCCCAACAGTAATAGAGAGTCAGCCACCAAAAGTAGTAATAGACCATTCGGGGGG 3946
DB |||||
QY 4140 TCAGTAGTCAAGAGCGCAACAAATTTTCAGTACAGGGAACCTTTTTCACATCTTCAGAA 4199
DB |||||
QY 3947 TCAGTAGTCAAGAGCGCAACAAATTTTCAGTACAGGGAACCTTTTTCAGATCTTCAGAA 4006
DB |||||
QY 4200 AGTTTGTATTCAGTAGTCAATTTGCGGAGCATCAATTAATGGGATTTATACAGAGCAACA 4259
DB |||||
QY 4007 AGTTTGTATTCAGTAGTCAATTTGCGGAGCATCAATTAATGGGATTTATACAGAGCAACA 4066
DB |||||
QY 4260 GTGGAGTCAACATCTTACCAACTTTTGGGCTCTCAGAAAGCATAAACAGTTCTACTACG 4319
DB |||||
QY 4067 GTGGAGTCAACATCTTACCAACTTTTGGGCTCTCAGAAAGCATAAACAGTTCTACTACG 4126
DB |||||
QY 4320 CCATTAGTGAAACTTTTTCAAATTCGCCCAGTGGAGAGAAAGGACACAGCGATACGCA 4379
DB |||||
QY 4127 CCATTAGTGAAACTTTTTCAAATTCGCCCAGTGGAGAGAAAGGACACAGCGATACGCA 4186
DB |||||
QY 4380 TTAGCGGCAAGGATGCAACTTTTATCAACAGGCTCTTATAGATTAACCTAGCGCTGG 4439
DB |||||
QY 4187 TTAGCGGCAAGGATGCAACTTTTATCAACAGGCTCTTATAGATTAACCTAGCGCTGG 4246
DB |||||
QY 4440 ATCATCTTTTGGACAACTTTTCTTCCAAATCTAGCTCCAAATCACTTTCTTATAGTACCA 4499
DB |||||
QY 4247 ATCATCTTTTGGACAACTTTTCTTCCAAATCTAGCTCCAAATCACTTTCTTATAGTACCA 4306
DB |||||
QY 4500 TTATTGTCAACTTGCAGAACTTTCATAGCTCTCAATTTGGTCTCTGTAAACGAT 4559
DB |||||
QY 4307 TTATTGTCAACTTGCAGAACTTTCATAGCTCTCAATTTGGTCTCTGTAAACGAT 4366
DB |||||
QY 4560 GACTCAACTTGCAGAACTTAACTTGAAGCTCAGTTCAGTGAACCTTCAAGGTTGTGC 4619
DB |||||
QY 4367 GACTCAACTTGCAGAACTTAACTTGAAGCTCAGTTCAGTGAACCTTCAAGGTTGTGC 4426
DB |||||
QY 4620 AGCTGCTCAGCAGCATAGGAAACACGCTTTTCTTACCAAACTCAAGGAATTTCAAC 4679
DB |||||
QY 4427 AGCTGCTCAGCAGCATAGGAAACACGCTTTTCTTACCAAACTCAAGGAATTTCAAC 4486
DB |||||
QY 4680 TCTGCACACTTGGCTATGTCAGGTAGCAAGGAAATGTCTACTTTGAAGTCCGACAGTGA 4739
DB |||||

QY 5807 GTACGCTGATTTTAAACGTGAATTTATCTCAAGATCTCTGCTCGCGCGTTTCGGTGAT 5866
Db 6705 GTACGCTGATTTTAAACGTGAATTTATCTCAAGATCTCTGCTCGCGCGTTTCGGTGAT 6764
QY 5867 CACGGTGAACCTCTGACACATGACGCTCCCGGAGACGGTCAACAGCTTGTCTGAAGCG 5926
Db 6765 CACGGTGAACCTCTGACACATGACGCTCCCGGAGACGGTCAACAGCTTGTCTGAAGCG 6824
QY 5927 GATGCGGAGGACACAGCCCGTCAAGGCGCGTCAAGGCGTGTGCGGGTGTGCGGGC 5986
Db 6825 GATGCGGAGGACACAGCCCGTCAAGGCGCGTCAAGGCGTGTGCGGGTGTGCGGGC 6884
QY 5987 GCAGCCATGACCCAGTACGCTAGCGATAGCGGAGTGTATCTGGCTTAACTATCGGCGAT 6046
Db 6885 GCAGCCATGACCCAGTACGCTAGCGATAGCGGAGTGTATCTGGCTTAACTATCGGCGAT 6944
QY 6047 CAGACGAGTTGTAAGTGAAGTGCACATATGCGGTGTGAATACCGACAGATGCGTAA 6106
Db 6945 CAGACGAGTTGTAAGTGAAGTGCACATATGCGGTGTGAATACCGACAGATGCGTAA 7004
QY 6107 GGAGAAATACCGCATCAGGCGCTCTCCGCTTCTCGCTCACTGACTCGCTGCGCTCG 6166
Db 7005 GGAGAAATACCGCATCAGGCGCTCTCCGCTTCTCGCTCACTGACTCGCTGCGCTCG 7064
QY 6167 TCGTTCGGCTCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAACTACGGTTATCCACAG 6226
Db 7065 TCGTTCGGCTCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAACTACGGTTATCCACAG 7124
QY 6227 AATCAGGGGTAACGACGAAAGACATGTGACGAAAGCCAGCAAGCCAGCAAGC 6286
Db 7125 AATCAGGGGTAACGACGAAAGACATGTGACGAAAGCCAGCAAGCCAGCAAGC 7184
QY 6287 GTAAAAAGCGCGGTGCTGCGCTTTTCCATAGGCTCCGCGCGCTTACAGGATCACA 6346
Db 7185 GTAAAAAGCGCGGTGCTGCGCTTTTCCATAGGCTCCGCGCGCTTACAGGATCACA 7244
QY 6347 AATATCGAGCTCAAGTCAGAGTGGGGAACCCGACAGCACTATAAGATACAGGCGT 6406
Db 7245 AATATCGAGCTCAAGTCAGAGTGGGGAACCCGACAGCACTATAAGATACAGGCGT 7304
QY 6407 TTCCCGCTTGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCCTGCGCTTACCGGATACC 6466
Db 7305 TTCCCGCTTGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCCTGCGCTTACCGGATACC 7364
QY 6467 TGTCCGCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCACTAGCTCACTGCTAGGTATC 6526
Db 7365 TGTCCGCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCACTAGCTCACTGCTAGGTATC 7424
QY 6527 TCAGTTCGGTGTAGGTGCTGCTCCAGCTGGGCTGTGCGAGACCCCGCTTACG 6586
Db 7425 TCAGTTCGGTGTAGGTGCTGCTCCAGCTGGGCTGTGCGAGACCCCGCTTACG 7484
QY 6587 CCGACCGCTGCGCTTATCCGGTAACATCGTCTTGAGTCCAAACCCCGGTAAAGCACT 6646
Db 7485 CCGACCGCTGCGCTTATCCGGTAACATCGTCTTGAGTCCAAACCCCGGTAAAGCACT 7544
QY 6647 TATCGCCATCGGACGACCACTGGTAACAGGATTACAGAGCGAGTATGTAGGCGGTG 6706
Db 7545 TATCGCCATCGGACGACCACTGGTAACAGGATTACAGAGCGAGTATGTAGGCGGTG 7604
QY 6707 CTACAGAGTTCTTGAAGTGTGGCTTAACCTACGCTTACCTAGAGGACAGTATTTGGTA 6766
Db 7605 CTACAGAGTTCTTGAAGTGTGGCTTAACCTAGAGGACAGTATTTGGTA 7664
QY 6767 TCTCGGCTCTGCTGAAGCGAGTACCTTCGGAAGAGTGTGAGCTCTTGATCCGGCA 6826
Db 7665 TCTCGGCTCTGCTGAAGCGAGTACCTTCGGAAGAGTGTGAGCTCTTGATCCGGCA 7724
QY 6827 AACAAACCCAGCGCTGAGCGGTGGTGTGTTTGTGCAAGCAGCAGATACCGCGAGAA 6886
Db 7725 AACAAACCCAGCGCTGAGCGGTGGTGTGTTTGTGCAAGCAGCAGATACCGCGAGAA 7784

QY 6887 AAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGCTCTGACGCTCAGTGAACG 6946
Db 7785 AAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGCTCTGACGCTCAGTGAACG 7844
QY 6947 AAAACTCAGTTAAGGATTTTGGTCAAGATTAATCAAAAAGGATCTTCACTAGATCC 7006
Db 7845 AAAACTCAGTTAAGGATTTTGGTCAAGATTAATCAAAAAGGATCTTCACTAGATCC 7904
QY 7007 TTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATAGTAATAACTTGGTCTG 7066
Db 7905 TTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATAGTAATAACTTGGTCTG 7964
QY 7067 ACAGTTACCAATGCTTAATCAGTAGGAGCACCTATCTCAGCGATCTGTCTATTGGTTCA 7126
Db 7965 ACAGTTACCAATGCTTAATCAGTAGGAGCACCTATCTCAGCGATCTGTCTATTGGTTCA 8024
QY 7127 CCATAGTTGCTGACTCCCGCTGCTGTAGATAAATCACTACGATACGGAGGCTTACCATCTG 7186
Db 8025 CCATAGTTGCTGACTCCCGCTGCTGTAGATAAATCACTACGATACGGAGGCTTACCATCTG 8084
QY 7187 GCCCAGTGTCTGCAATGATACCGCGAGACCCACGCTCACCGCTCCAGATTTATCAGCAA 7246
Db 8085 GCCCAGTGTCTGCAATGATACCGCGAGACCCACGCTCACCGCTCCAGATTTATCAGCAA 8144
QY 7247 TAAACACGACGAGCCGGAAGGCGAGCGAGAGTGTCTGCAACTTTTATCCGCTCCA 7306
Db 8145 TAAACACGACGAGCCGGAAGGCGAGCGAGAGTGTCTGCAACTTTTATCCGCTCCA 8204
QY 7307 TCCAGTCTATTAAATTTGCGGGAAGCTAGAGTAAGTGTTCGCCAGTAAATAGTTTCG 7366
Db 8205 TCCAGTCTATTAAATTTGCGGGAAGCTAGAGTAAGTGTTCGCCAGTAAATAGTTTCG 8264
QY 7367 GCAAGCTGTTGCAATCTGTCAGCATCTGCTGTCACGCTCGCTGCTGCTGATGCTT 7426
Db 8265 GCAAGCTGTTGCAATCTGTCAGCATCTGCTGTCACGCTCGCTGCTGCTGATGCTT 8324
QY 7427 CATTCAGCTCCGTTCCCAACGATCAGCGCAGTACATGATCCCCCATGTTGCAAAA 7486
Db 8325 CATTCAGCTCCGTTCCCAACGATCAGCGCAGTACATGATCCCCCATGTTGCAAAA 8384
QY 7487 AAGCGTTAGCTCTCTCGCTCTCCGATCTGTCAGAAAGTAAAGTTGGCCGAGTGTAT 7546
Db 8385 AAGCGTTAGCTCTCTCGCTCTCCGATCTGTCAGAAAGTAAAGTTGGCCGAGTGTAT 8444
QY 7547 CACTCATGTTATGCGCAGCTGCAATTTCTCTTACTGTCTATGCCATCCCTAGATGCT 7606
Db 8445 CACTCATGTTATGCGCAGCTGCAATTTCTCTTACTGTCTATGCCATCCCTAGATGCT 8504
QY 7607 TTTCTGTGACTGGTGAAGTCAACCAAGTCAATCTGAGAAATAGTGTATGCGGACCGA 7666
Db 8505 TTTCTGTGACTGGTGAAGTCAACCAAGTCAATCTGAGAAATAGTGTATGCGGACCGA 8564
QY 7667 GTTGTCTTTCGCGCGTCAACAGGGATAATACCGGCCACATAGCAGAACTTTAAAG 7726
Db 8565 GTTGTCTTTCGCGCGTCAACAGGGATAATACCGGCCACATAGCAGAACTTTAAAG 8624
QY 7727 TGCTCATCATTTGGAAGAGTCTTTCGCGGCGAAACTCTCAAGGATCTTACCGCTGTGA 7786
Db 8625 TGCTCATCATTTGGAAGAGTCTTTCGCGGCGAAACTCTCAAGGATCTTACCGCTGTGA 8684
QY 7787 GATCAGTTTCGATGTAAACCACTCGTGCACCAACTGATCTTCAGCATCTTTTACTTCA 7846
Db 8685 GATCAGTTTCGATGTAAACCACTCGTGCACCAACTGATCTTCAGCATCTTTTACTTCA 8744
QY 7847 CCAGGTTTCTGGTGAGCAAAAACAGAGCGCAAAATCCGCAAAAAGGATTAAGG 7906
Db 8745 CCAGGTTTCTGGTGAGCAAAAACAGAGCGCAAAATCCGCAAAAAGGATTAAGG 8804
QY 7907 CCACGCAAAATGTTGAATCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7966
Db 8805 CCACGCAAAATGTTGAATCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8864
QY 7967 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 8026

Db 3661 CAGAAATCACCTGGGAATCAATACCATGTTGAGTTCAGTTCAGGAGGTTCTGAGGCAACGAA 3719
Qy 2744 ATCTGGATCAGCGTATTTATCAGCAATTAATAGAACTTCAGAGGCCAGCAGCATGTC 2803
Db 3720 ATCTGGATCAGCGTATTTATCAGCAATTAATAGAACTTCAGAGGCCAGCAGCATGTC 3779
Qy 2804 AATACTACACAGGCGTATGATGTCATTTGAAACCATCATCTTGGCAGCAGTAACGAACTG 2863
Db 3780 AATACTACACAGGCGTATGATGTCATTTGAAACCATCATCTTGGCAGCAGTAACGAACTG 3839
Qy 2864 GTTTCCTGGACCAATATTTTGTGCACACTTAGGAACAGTTTCTGTTCCGTAAGCCATAGC 2923
Db 3840 GTTTCCTGGACCAATATTTTGTGCACACTTAGGAACAGTTTCTGTTCCGTAAGCCATAGC 3999
Qy 2924 AGCTACTCCCTGGCGCTCTGCTAGCACGATACACTTAGCACCAACCTTGTGGGCAAC 2983
Db 3900 AGCTACTCCCTGGCGCTCTGCTAGCACGATACACTTAGCACCAACCTTGTGGGCAAC 3959
Qy 2984 GTAGATGACTTCTGGGGTAAGGGTACCATTCCTTTAGTGGAGATGCAAAAACAATTC 3043
Db 3960 GTAGATGACTTCTGGGGTAAGGGTACCATTCCTTTAGTGGAGATGCAAAAACAATTC 4019
Qy 3044 TTTGCAACCCAGCACTTTGGCAGGACACCCAGCATCAGGAGTGGAGGCAAGTTC 3103
Db 4020 TTTGCAACCCAGCACTTTGGCAGGACACCCAGCATCAGGAGTGGAGGCAAGTTC 4079
Qy 3104 GGTTCACACGAGAAATATAGAGCCAACTTCTCAATAGTCTTGCACAAACAGAGCAGAC 3163
Db 4080 GGTTCACACGAGAAATATAGAGCCAACTTCTCAATAGTCTTGCACAAACAGAGCAGAC 4139
Qy 3164 TACACAGGCGAGTCTCAACTGTCAGCACTGCTCGTTAGTTGAGTTTCATGGATTTCCCT 3223
Db 4140 TACACAGGCGAGTCTCAACTGTCAGCACTGCTCGTTAGTTGAGTTTCATGGATTTCCCT 4199
Qy 3224 GAGGTATCTATAGAGAGATCAATGGCTCTCTTAACGTTATCTGGCAATTCATAAGTTC 3283
Db 4200 GAGGTATCTATAGAGAGATCAATGGCTCTCTTAACGTTATCTGGCAATTCATAAGTTC 4259
Qy 3284 CTCTGGGAAAGAGCTTCTAACACAGGTGCTTCAAAGCGACTCCATCAAACTTTGGCAGT 3343
Db 4260 CTCTGGGAAAGAGCTTCTAACACAGGTGCTTCAAAGCGACTCCATCAAACTTTGGCAGT 4319
Qy 3344 TAGTTCTAAAGGGCTTGTACCACTTTGAGACATTTGCAAACTGTTGACTTAA 3403
Db 4320 TAGTTCTAAAGGGCTTGTACCACTTTGAGACATTTGCAAACTGTTGACTTAA 4379
Qy 3404 TTCCATAATCTGTTCCGTTTCTGGATAGGACGACGAGGCACTTCAATTTCTTTGTA 3463
Db 4380 TTCCATAATCTGTTCCGTTTCTGGATAGGACGACGAGGCACTTCAATTTCTTTGTA 4439
Qy 3464 GGAGGCTTAGAAAGCTCAATTTGCACTTCAATACGACCTTCAAGGGACTTCTTT 3523
Db 4440 GGAGGCTTAGAAAGCTCAATTTGCACTTCAATACGACCTTCAAGGGACTTCTTT 4499
Qy 3524 AGGTTTGGATTTCTTTTAGTGTGTTCTTGGTGTATCTGGCTTGGCATCTCTCTTCTCT 3583
Db 4500 AGGTTTGGATTTCTTTTAGTGTGTTCTTGGTGTATCTGGCTTGGCATCTCTCTTCTCT 4559
Qy 3584 TCTAGTGACCTTAGGACTTCATATCCAGTTTCTCTCCACTCGTCCAGCTCACACC 3643
Db 4560 TCTAGTGACCTTAGGACTTCATATCCAGTTTCTCTCCACTCGTCCAGCTCACACC 4619
Qy 3644 GTACTTGGCAGACTTAACTAATGCAAAATAAATAGTCAGACATTCACGCTATATC 3703
Db 4620 GTACTTGGCAGACTTAACTAATGCAAAATAAATAGTCAGACATTCACGCTATATC 4679
Qy 3704 TTCTTTGGATTTAGCTTCTGCAAGTTTCATCAGCTTCTCTCCCTTAATTTTAGCGTTTAC 3762
Db 4680 TTCTTTGGATTTAGCTTCTGCAAGTTTCATCAGCTTCTCTCCCTTAATTTTAGCGTTTAC 4739
Qy 3763 AAATCTGCTGCAAAATAACCGTTTGGTATAGAACCTTCTGGAGATTTGCTTTACGAT 3822

Db 4740 AAACCTTCGTCGTCATAAATAACCGTTTGGTATTAAGAACCTTCTGGAGCATTTGCTCTTACGAT 4799
Qy 3823 GCCACAAGGTGGCTTCCATGGCTCTAAGAACCTTGTATTTGGCCAAAACAGGAAGTCCGTT 3882
Db 4800 CCCACAGGT-CTTCCATGGCTCTAAGAACCTTGTATTTGGCCAAAACAGGAAGTCCGTT 4858
Qy 3883 CCAAGTGACAGAAAACCAACCTGTTGTTCAACACAAATTTCAAGCAGTCTCCATCAC 3942
Db 4859 CCAAGTGACAGAAAACCAACCTGTTGTTCAACACAAATTTCAAGCAGTCTCCATCAC 4918
Qy 3943 AATCCAAATTCGATACCCAGCAACTTTTGCAGTTGCTCCAGATGTAGCACCTTTATACCACA 4002
Db 4919 AATCCAAATTCGATACCCAGCAACTTTTGCAGTTGCTCCAGATGTAGCACCTTTATACCACA 4978
Qy 4003 AACCGTGACAGAGATTTGGTGTAGCTCAGTTTGTGCTTATAGCTCCGGAATAGACT 4062
Db 4979 AACCGTGACAGAGATTTGGTGTAGCTCAGTTTGTGCTTATAGCTCCGGAATAGACT 5038
Qy 4063 TTTTGGACAGATACACAGGCCCAACGAGTAAATTAAGAGTTCAGCCACCAAGTAGTGA 4122
Db 5039 TTTTGGACAGATACACAGGCCCAACGAGTAAATTAAGAGTTCAGCCACCAAGTAGTGA 5098
Qy 4123 ATAGACCATCGGGGGCGGTCAAGTAGTCAAAAGACGCAACAAATTTTCACTGACAGGGAAT 4182
Db 5099 ATAGACCATCGGGGGCGGTCAAGTAGTCAAAAGACGCAACAAATTTTCACTGACAGGGAAT 5158
Qy 4183 TTTTGTAGCTTTTCCAGAAAGTTTGTATTTAGTGTCAATTTGCCAGCATCAATAATGGGGA 4242
Db 5159 TTTTGTAGCTTTTCCAGAAAGTTTGTATTTAGTGTCAATTTGCCAGCATCAATAATGGGGA 5218
Qy 4243 TTATACAGAGCAACAGTGGAGTCAACATCTACCAACTTTGCGGTCTCAGAAAAAGCAT 4302
Db 5219 TTATACAGAGCAACAGTGGAGTCAACATCTACCAACTTTGCGGTCTCAGAAAAAGCAT 5278
Qy 4303 AAACAGTTTCTATACCGCATTTAGTGAAACTTTTCAATCGCCAGTGGAGGAAAG 4362
Db 5279 AAACAGTTTCTATACCGCATTTAGTGAAACTTTTCAATCGCCAGTGGAGGAAAG 5338
Qy 4363 GCACAGCATCTAGCATTAGCGGCAAGGATCAAACTTTATCAACAGGCTCTATAGA 4422
Db 5339 GCACAGCATCTAGCATTAGCGGCAAGGATCAAACTTTATCAACAGGCTCTATAGA 5398
Qy 4423 TAAACCTTAGCGCTGGGATCATCTTTTGGACAACTCTTTCTGCCAAATCTAGTCCAAA 4482
Db 5399 TAAACCTTAGCGCTGGGATCATCTTTTGGACAACTCTTTCTGCCAAATCTAGTCCAAA 5458
Qy 4483 TCACCTTATGATACCAATTTATGACAACTTGGACAACTTGTGATCAGCTCTCAAT 4542
Db 5459 TCACCTTATGATACCAATTTATGACAACTTGGACAACTTGTGATCAGCTCTCAAT 5518
Qy 4543 GGTCTCTGTAAACGGATGACTCAACTTGCATTAACCTTGAAGCTCAGTCAATTTAGTGA 4602
Db 5519 GGTCTCTGTAAACGGATGACTCAACTTGCATTAACCTTGAAGCTCAGTCAATTTAGTGA 5578
Qy 4603 ACTTGATCAGGTTGTGAGCTGTGTCAGCAGCATAGGAAACAGGCTTTTCTTACCAAC 4662
Db 5579 ACTTGATCAGGTTGTGAGCTGTGTCAGCAGCATAGGAAACAGGCTTTTCTTACCAAC 5638
Qy 4663 TCAGGAAATTTATCAAACTCTGCAACACTTGCATTCAGGTATGCAAGGAAATGTCTATC 4722
Db 5639 TCAGGAAATTTATCAAACTCTGCAACACTTGCATTCAGGTATGCAAGGAAATGTCTATC 5698
Qy 4723 TTGAAGTCGACAGTGTAGTGTGTTGAGAAATTTCTGAAGCGGTATTTTATTTATCAGT 4782
Db 5699 TTGAAGTCGACAGTGTAGTGTGTTGAGAAATTTCTGAAGCGGTATTTTATTTATCAGT 5758
Qy 4783 GAGTCAGTCTATCAGGAGATCTCTTACGCGGACGATCTGTCGCGGCTCAGGTCGGCA 4842
Db 5759 GAGTCAGTCTATCAGGAGATCTCTTACGCGGACGATCTGTCGCGGCTCAGGTCGGCA 5818
Qy 4843 TCACCGGCGGACAGGTCGCTGCTGGCGCTTATATCGCCGACATCACCGATGGGGAAG 4902
Db 5819 TCACCGGCGGACAGGTCGCTGCTGGCGCTTATATCGCCGACATCACCGATGGGGAAG 5878

QY 4903 ATCCGGCTCGCCACTTCGGGCTCATGACGCTTCTTTTCGGCGTGGGTATGTTGGCAGGCC 4962
DB 5879 ATCCGGCTCGCCACTTCGGGCTCATGAGCGCTTCTTTTCGGCGTGGGTATGTTGGCAGGCC 5938
QY 4963 CCGTGGCCGGGGAGCTGTGGGCGCCATCTCTTTCGATGAC 5004
DB 5939 CCGTGGCCGGGGAGCTGTGGGCGCCATCTCTTTCGACCTGCAGGGGGGGGGGGAAG 5998
QY 5005 ----- 5004
DB 5999 CCACGTTGTCTCAAAATCTCTGATGTATACATTGCAACAGATAAAATATATCATCATG 6058
QY 5005 ----- 5004
DB 6059 AACATAAAAGCTGTCTGTACATAAACAGTAATACAGGGGTGTATGAGCCATATCA 6118
QY 5005 ----- 5004
DB 6119 ACGGGAACGCTTCTCTCAAGCCGCGATTAATTCACATGATGCTGATTAATG 6178
QY 5005 ----- 5004
DB 6179 GTATAAATGGCTCGCGATAATGTCCGGCAATCAGGTGCGACAATCTATCGATTGTATGG 6238
QY 5005 ----- 5004
DB 6239 GAAGCCGATGGCGAGAGTGTCTTCTGAACATGCGAAAGGTAGCGTTCGCAATGATGT 6298
QY 5005 ----- 5004
DB 6299 TACAGATGAGATGTCAGACTAACTGCTGACGGATTTATGCTCTTCGACCATCAA 6358
QY 5005 ----- 5004
DB 6359 GCATTTATCGGTACTCCTGATGATGATGCTTACTCACCACCTGCGATCCCGGAAAC 6418
QY 5005 ----- 5004
DB 6419 ASCATTCCAGTATTAGAGAAATATCCTGATTCAGGTGAAATATTTGTTGATGGCTGGC 6478
QY 5005 ----- 5004
DB 6479 AGTGTCTCGCGGTGTCATTCGATTCGATTCCTGTTTGTAAATGTCCTTTTAAACAGCGATCG 6538
QY 5005 ----- 5004
DB 6539 CGTATTCGTCTCGCTCAGCGCAATCAGCAATGAATAACGGTTTGTGATCGGAGTGA 6598
QY 5005 ----- 5004
DB 6599 TTTTGATCAGCAGGTAATGCTGCGCTGTTGACAAAGTCTGGAAGAAATGCAATAGCT 6658
QY 5005 ----- 5004
DB 6659 TTTGCCATTCACCGGATTCAGTCGTCACTCATGTTGATTTCTCACTTGATAACCTTAT 6718
QY 5005 ----- 5004
DB 6719 TTTTACGAGGGAAATTAATAGTTGTATTGATGTTGGACGAGTCGGAATCGACACCG 6778
QY 5005 ----- 5004
DB 6779 ATACCAGGATCTTGCCATCTTATGAAATGATTAATCCTGATATGAATAATGTCAGTTTCA 6838
QY 5005 ----- 5004
DB 6839 ACGCTTTTCAAAATATGTTATGATTAATCCTGATATGAATAATGTCAGTTTCA 6898
QY 5005 ----- 5004
DB 6899 GATGCTCGATGATTTTCTTAATCAGAAATGGTTAAATGGTTGTTGTAACATCGCGAGAGCAT 6958

QY 5005 ----- 5004
DB 6959 TACGCTGACTTGACGGGACGGCGCTTTGTTGAATAAATCGAACTTTTGTGCTGAGTTGAAG 7018
QY 5005 ----- 5004
DB 7019 GATCAGATCACGCACTTCTCCCGACAACCGACACCGTCCGTGGCAAGCAAAAGTTCAAA 7078
QY 5005 ----- 5004
DB 7079 ATCACCAACTGCTCCACCTACAAACAAGCTCTCATCAACCGTGGCTCCCTCACCTTCTGG 7138
QY 5005 ----- 5004
DB 7139 CTGSGATGATGGGGGATTCAGGCTTGATGAGTCAGCAACACCTTCTTACGAGGCGAGA 7198
QY 5005 ----- 5042
DB 7199 CCTCAGCCGCCCCCTGAGGTCCACGGCGCGTCTCAACGGCTCAACCTA 7258
QY 5043 CTACTGGGCTGCTTCTTAATGAGAGTCGCATTAAGGAGAGCGTCGAGTATCTATGATT 5102
DB 7259 CTACTGGGCTGCTTCTTAATGAGAGTCGCATTAAGGAGAGCGTCGAGTATCTATGATT 7318
QY 5103 GGAAGTATGGGAATGGTGATACCCGCAATTCCTCAGTGTCTTGAAGTCTCTCTATCAGATTA 5162
DB 7319 GGAAGTATGGGAATGGTGATACCCGCAATTCCTCAGTGTCTTGAAGTCTCTCTATCAGATTA 7378
QY 5163 TGCCCAACTAAAGCAACCGGAGGAGATTCATGTAATTTCTCTGACTTTGGTCA 5222
DB 7379 TGCCCAACTAAAGCAACCGGAGGAGATTCATGTAATTTCTCTGACTTTGGTCA 7438
QY 5223 TCAGTAGACTCGAAGTGTGAGACTATCTCGGTATGACAGCAGAAATCTCTCTCTCGAG 5282
DB 7439 TCAGTAGACTCGAAGTGTGAGACTATCTCGGTATGACAGCAGAAATCTCTCTCTCGAG 7498
QY 5283 ACAGTAATGAAGTCCCAACCAATTAAGAAATCCTTGTATCAGGAACAACCTCTTGT 5342
DB 7499 ACAGTAATGAAGTCCCAACCAATTAAGAAATCCTTGTATCAGGAACAACCTCTTGT 7558
QY 5343 CGAACTTTTGGTGGCTTGAATATAAATGATAGAGTGGATATGTCGGGTAGCAATGGA 5402
DB 7559 CGAACTTTTGGTGGCTTGAATATAAATGATAGAGTGGATATGTCGGGTAGCAATGGA 7618
QY 5403 GCGGCAATGCTTACCTTCTGGACCTTCAAGAGTATGTAGGTTGTAGATCTCATG 5462
DB 7619 GCGGCAATGCTTACCTTCTGGACCTTCAAGAGTATGTAGGTTGTAGATCTCATG 7678
QY 5463 CGAACTTCAAGTGAACCTTGTCTTTCGTTCAACCAATTCGGAATCCAGAAATCAAA 5522
DB 7679 CGAACTTCAAGTGAACCTTGTCTTTCGTTCAACCAATTCGGAATCCAGAAATCAAA 7738
QY 5523 GTTGTCTCTACTATTGATCCAGCCAGTCGGTCTTGAACCTGACAATAGTGTGTCG 5582
DB 7739 GTTGTCTCTACTATTGATCCAGCCAGTCGGTCTTGAACCTGACAATAGTGTGTCG 7798
QY 5583 TGTGTGTAGGTCATCTTTGTATGAATAAATCTAGTCTTTGATCTAAATAATCTTTGACGAG 5642
DB 7799 TGTGTGTAGGTCATCTTTGTATGAATAAATCTAGTCTTTGATCTAAATAATCTTTGACGAG 7858
QY 5643 CCAAGCCGCAATTAACCCAAATCTAAACTCTTTTAAACGTTAAAGGACAAGTATGTC 5702
DB 7859 CCAAGCCGCAATTAACCCAAATCTAAACTCTTTTAAACGTTAAAGGACAAGTATGTC 7918
QY 5703 TGCCGTATTTAAACCCCAATCAGCTCGTGTCTCATCTCATCACTTGAGGGGCACTA 5762
DB 7919 TGCCGTATTTAAACCCCAATCAGCTCGTGTCTCATCTCATCACTTGAGGGGCACTA 7978
QY 5763 TCTGTGTAGAGAAATTTGGGAGATCGGATATCGAGAAAGGTACGCTGATTTTAA 5822
DB 7979 TCTGTGTAGAGAAATTTGGGAGATCGGATATCGAGAAAGGTACGCTGATTTTAA 8038
QY 5823 CGTGAATTTTATCTCAAGATCTCTGCTCGCGCGTTCGGTGTGATGACGGTGAACCTCT 5882

Db 8039 CGTGAATTTATCTCAAGATCTCGCTCGCGGTTTCGGTGATGACGGTGAAACCTCT 8098
QY 5883 GACATGAGCTCCCGAGAGCGTCAAGCTTGTCTGTAAAGCGATGCCGGAGCAGAC 5942
Db 8099 GACATGAGCTCCCGAGAGCGTCAAGCTTGTCTGTAAAGCGATGCCGGAGCAGAC 8158
QY 5943 AAGCCCGTCAGGCGCGCTCAGCGGCTGTTGGGGGTGTCGGGGCGACGCCATGACCCAGT 6002
Db 8159 AAGCCCGTCAGGCGCGCTCAGCGGCTGTTGGGGGTGTCGGGGCGACGCCATGACCCAGT 8218
QY 6003 CACGTAGCGATAGCGAGGTGATATCTGCTTAACTATCGGCATCAGAGCAGATTGACT 6062
Db 8219 CACGTAGCGATAGCGAGGTGATATCTGCTTAACTATCGGCATCAGAGCAGATTGACT 8278
QY 6063 GAGGTGACCATATGCGGTGTAATACCGCACAGATGCGTAAGGAGAAATACCGCAT 6122
Db 8279 GAGGTGACCATATGCGGTGTAATACCGCACAGATGCGTAAGGAGAAATACCGCAT 8338
QY 6123 CAGGCGCTCTTCGGCTTCTCGCTCACTGACTCGCTGCGCTCGGTTCGGCTGCGGCG 6182
Db 8339 CAGGCGCTCTTCGGCTTCTCGCTCACTGACTCGCTGCGCTCGGTTCGGCTGCGGCG 8398
QY 6183 AGCGTATCAGCTACTCAAGCGGTAAATAGCGTTATCCAGATCAGGGGATACGC 6242
Db 8399 AGCGTATCAGCTACTCAAGCGGTAAATAGCGTTATCCAGATCAGGGGATACGC 8458
QY 6243 AGAAAGAACATGTCAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAGAGGCCGCTT 6302
Db 8459 AGAAAGAACATGTCAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAGAGGCCGCTT 8518
QY 6303 GCTGCGCTTTTCCATAGGCTCCGCCCTTGAAGAGCATCACAAAATCAAGCTCAAG 6362
Db 8519 GCTGCGCTTTTCCATAGGCTCCGCCCTTGAAGAGCATCACAAAATCAAGCTCAAG 8578
QY 6363 TCAGAGGTGGGAAACCCGACGAGACTATAAGATACCGAGGTTCCCGCTGGAAGCTC 6422
Db 8579 TCAGAGGTGGGAAACCCGACGAGACTATAAGATACCGAGGTTCCCGCTGGAAGCTC 8638
QY 6423 CCTGCTGGCTCTCTGTTCCGACCCCTGCGGTTTACCGGATACCTGTCCGCTTCTCCC 6482
Db 8639 CCTGCTGGCTCTCTGTTCCGACCCCTGCGGTTTACCGGATACCTGTCCGCTTCTCCC 8698
QY 6483 TTCGGGAAGCGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGT 6542
Db 8699 TTCGGGAAGCGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGT 8758
QY 6543 CGTTCGCTCCAGCTGGGCTGTGCAAGACCCCGCTTCAAGCCGACCGCTGCGCTT 6602
Db 8759 CGTTCGCTCCAGCTGGGCTGTGCAAGACCCCGCTTCAAGCCGACCGCTGCGCTT 8818
QY 6603 ATCCGGTAACCTATCGTCTGTAGTCCAAACCCCGTAAAGACACGACTTATCGCCACTGGCAGC 6662
Db 8819 ATCCGGTAACCTATCGTCTGTAGTCCAAACCCCGTAAAGACACGACTTATCGCCACTGGCAGC 8878
QY 6663 AGCCACTGTAACAGGATTCAGAGCGAGGTATGAGCGGTGCTACAGAGTCTTGAA 6722
Db 8879 AGCCACTGTAACAGGATTCAGAGCGAGGTATGAGCGGTGCTACAGAGTCTTGAA 8938
QY 6723 GTGTGGCTTAACCTACCGCTACACTAGAGGACAGTATTTGGTATCTCGCTGCTGCTGAA 6782
Db 8939 GTGTGGCTTAACCTACCGCTACACTAGAGGACAGTATTTGGTATCTCGCTGCTGCTGAA 8998
QY 6783 GCCAGTTACCTTCGAAAGAGTTGAGTCTTGTATCCGCAACCAACCCGCTGG 6842
Db 8999 GCCAGTTACCTTCGAAAGAGTTGAGTCTTGTATCCGCAACCAACCCGCTGG 9058
QY 6843 TAGCGGTGTTTTTTTGTGCAAGCAGCATACCGGCAAGAAAGAGTCTCAAGA 6902
Db 9059 TAGCGGTGTTTTTTTGTGCAAGCAGCATACCGGCAAGAAAGAGTCTCAAGA 9118
QY 6903 AGATCCTTTGATCTTTTCTACGGGCTTGAAGCTCAGTGGAAACGAAACTCAGCTTAAG 6962

Db 9119 AGATCCTTTGATCTTTTCTACGGGCTGACGCTCAGTGAACGAAACCTCAGCTTAAGG 9178
QY 5963 GATTTTGGTCATGAGATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATTTAAAAATG 7022
Db 9179 GATTTTGGTCATGAGATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATTTAAAAATG 9238
QY 7023 AAGTTTTAAATCAATCTAAAGTATATAGTAAACTTGGTCTGACAGTTTACCAATGCTT 7082
Db 9239 AAGTTTTAAATCAATCTAAAGTATATAGTAAACTTGGTCTGACAGTTTACCAATGCTT 9298
QY 7083 AATCAGTAGGACCTATCTCAGCGATCTGTCTATTTTGGTTCATCATAGTTGCTGACT 7142
Db 9299 AATCAGTAGGACCTATCTCAGCGATCTGTCTATTTTGGTTCATCATAGTTGCTGACT 9358
QY 7143 CCCGCTCGTGTAGATACCTACGATACGAGGCTTACCATCTGGCCCCAGAGTCTCAAT 7202
Db 9359 CCCGCTCGTGTAGATACCTACGATACGAGGCTTACCATCTGGCCCCAGAGTCTCAAT 9418
QY 7203 GATACCGGAGACCCAGCTCCAGCTCCAGATTTATCAGCAATTAACCCAGCCAGCCGG 7262
Db 9419 GATACCGGAGACCCAGCTCCAGCTCCAGATTTATCAGCAATTAACCCAGCCAGCCGG 9478
QY 7263 AAGGCGGAGCGCAGAGTGGTCTCGAACTTATCCGCTCCATCCAGTCTATTAATTTG 7322
Db 9479 AAGGCGGAGCGCAGAGTGGTCTCGAACTTATCCGCTCCATCCAGTCTATTAATTTG 9538
QY 7323 TTGCGCGGAAGCTAGAGTAAGTAGTTTGGCAGTAAATAGTTTGGCAAGAGTTTGGCCAT 7382
Db 9539 TTGCGCGGAAGCTAGAGTAAGTAGTTTGGCAGTAAATAGTTTGGCAAGAGTTTGGCCAT 9598
QY 7383 TGCTGACGGCATCGTGGTGTACGCTCGCTTGGTATGGCTTCAATTCAGCTCCGGTTC 7442
Db 9599 TGCTGACGGCATCGTGGTGTACGCTCGCTTGGTATGGCTTCAATTCAGCTCCGGTTC 9658
QY 7443 CCAACGATCAAGCGAGTTACATGATCCCGCATTTGTGCAAAAAGCGTTAGTCTCTT 7502
Db 9659 CCAACGATCAAGCGAGTTACATGATCCCGCATTTGTGCAAAAAGCGTTAGTCTCTT 9718
QY 7503 CGTCTCCGATCGTTGTGAGAAAGTGGCGGAGTGTATATCACTCATGTTATGCG 7562
Db 9719 CGTCTCCGATCGTTGTGAGAAAGTGGCGGAGTGTATATCACTCATGTTATGCG 9778
QY 7563 AGCACTGCATAATCTTCTTACTGTCTATGCCATCCGTAAGATGCTTCTGTCGACTGCTGA 7622
Db 9779 AGCACTGCATAATCTTCTTACTGTCTATGCCATCCGTAAGATGCTTCTGTCGACTGCTGA 9838
QY 7623 GTACTCAACCAAGTCAATCTGAGATAGTGTATGCGGACCGAGTTGCTCTTGCCCGGC 7682
Db 9839 GTACTCAACCAAGTCAATCTGAGATAGTGTATGCGGACCGAGTTGCTCTTGCCCGGC 9898
QY 7683 GTCAACACGGGATAATACCGCGCCACATAGCAGAACTTTAAAGTCTCATCATTTGGA 7742
Db 9899 GTCAACACGGGATAATACCGCGCCACATAGCAGAACTTTAAAGTCTCATCATTTGGA 9958
QY 7743 ACGTTCTTGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCGATGTA 7802
Db 9959 ACGTTCTTGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCGATGTA 10018
QY 7803 ACCCACTCGTGCAACCACTGATCTTCAAGTCTTTTACCTTTTCCAGCGTTTCTGGTG 7862
Db 10019 ACCCACTCGTGCAACCACTGATCTTCAAGTCTTTTACCTTTTCCAGCGTTTCTGGTG 10078
QY 7863 AGCAAAAACAGGAAGCAAAATCCCGCAAAAAGGGAATAAGGGCCACACGGAATGTTG 7922
Db 10079 AGCAAAAACAGGAAGCAAAATCCCGCAAAAAGGGAATAAGGGCCACACGGAATGTTG 10138
QY 7923 AATACTCATCTCTCTCTTTTCAATTAATTTGAAGCAATTTATCAGGTTTATGCTCAT 7982
Db 10139 AATACTCATCTCTCTCTTTTCAATTAATTTGAAGCAATTTATCAGGTTTATGCTCAT 10198
QY 7983 GAGCGGATACATATTTGAATGATTTAGAAAAATAAACAAGAGGTTTCCCGGCACTT 8042
Db 10199 GAGCGGATACATATTTGAATGATTTAGAAAAATAAACAAGAGGTTTCCCGGCACTT 10258

8043 TCCCGGAAAGTGCACCTGACGCTCTAAGAAACCATTTATCATGACATTAACCTATTA 8102
 Db
 10259 TCCCGGAAAGTGCACCTGACGCTCTAAGAAACCATTTATCATGACATTAACCTATTA 10318
 Qy
 8103 AAATAGCGTATACAGAGGCGCTTCGCTCTCAAGATTAATTTCTCATGTTTTCACAGCTT 8162
 Db
 10319 AAATAGCGTATACAGAGGCGCTTCGCTCTCAAGATTAATTTCTCATGTTTTCACAGCTT 10378
 Qy
 8163 ATCATCGATAAGCTGACTCATGTTGGTATTGTGAAATAGACGAGATCGGGAACACTGAA 8222
 Db
 10379 ATCATCGATAAGCTGACTCATGTTGGTATTGTGAAATAGACGAGATCGGGAACACTGAA 10438
 Qy
 8223 AAATAACAGTTATTATTCG 8241
 Db
 10439 AAATAACAGTTATTATTCG 10457
 Qy
 RESULT 13
 AX417109 10462 bp DNA linear PAT 14-JUN-2002
 LOCUS
 DEFINITION Sequence 34 from Patent WO0198344.
 ACCESSION AX417109
 VERSION AX417109.1 GI:21449696
 KEYWORDS
 SOURCE unidentified
 ORGANISM unidentified
 unclassified.
 1
 REFERENCE
 AUTHORS Ling, L. E. and Sanicola-Nadel, M.
 TITLE Angiogenesis: modulating compositions and uses
 JOURNAL Patent: WO 0198344-A 34 27-DEC-2001;
 BIOGEN, INC. (US)
 FEATURES
 source Location/Qualifiers
 1..10462
 /organism="unidentified"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32644"
 /note="Plasmid pUB116"
 ORIGIN
 Query Match 69.3%; Score 5709.8; DB 6; Length 10462;
 Best Local Similarity 77.4%; Pred. No. 0;
 Matches 8096; Conservative 0; Mismatches 142; Indels 2221; Gaps 8;
 2 GATCTAACATCAAGAGCGAAGGTTGAATGAACCTTTTGGCCATCCGACATCCACAGG 61
 Db 1 GATCTAACATCAAGAGCGAAGGTTGAATGAACCTTTTGGCCATCCGACATCCACAGG 60
 Qy 62 TCCATTTCTCACATAAGTGCACAAACGCAACAGAGGGGATACACTAGCAGCAGACCGTT 121
 Db 61 TCCATTTCTCACATAAGTGCACAAACGCAACAGAGGGGATACACTAGCAGCAGACCGTT 120
 Qy 122 GCAACGAGGACCTCCACTCTCTCTCTCAACACCCACTTTTGGCCATCGAAGAAACCA 181
 Db 121 GCAACGAGGACCTCCACTCTCTCTCTCAACACCCACTTTTGGCCATCGAAGAAACCA 180
 Qy 182 GCCAGTTATTGGGCTTGATTTGGAGCTCGCTCATTCCAATTCCTTATTAGGCTACTAA 241
 Db 181 GCCAGTTATTGGGCTTGATTTGGAGCTCGCTCATTCCAATTCCTTATTAGGCTACTAA 240
 Qy 242 CACCATGACTTTATTAGGCTGTCTATCTGGCCCGCTGGCGAGGTTCACTTTGTTTAT 301
 Db 241 CACCATGACTTTATTAGGCTGTCTATCTGGCCCGCTGGCGAGGTTCACTTTGTTTAT 300
 Qy 302 TTCCGAATGCAACAGCTCCGATTAACACCGGAAACATCACTCCAGATGAGGCTTTCTGA 361
 Db 301 TTCCGAATGCAACAGCTCCGATTAACACCGGAAACATCACTCCAGATGAGGCTTTCTGA 360
 Qy 362 GTGTGGGTCAAAATAGTTTCATGTTCCCGAAATGGCCCGAAACCTGACAGTTTAAACGCTG 421
 Db 361 GTGTGGGTCAAAATAGTTTCATGTTCCCGAAATGGCCCGAAACCTGACAGTTTAAACGCTG 420

422 TCTTGGACCTAATATGACAAAGCGTGATCTCATCAAGATGAACCTAAGTTTGGTTCTG 481
 Db 421 TCTTGGACCTAATATGACAAAGCGTGATCTCATCAAGATGAACCTAAGTTTGGTTCTG 480
 Qy 482 TGAATGTAAACGCGCAGTTGGTCAAAAAGAAAATTCCAAAAGTCGCCATACCGTTTGTG 541
 Db 481 TGAATGTAAACGCGCAGTTGGTCAAAAAGAAAATTCCAAAAGTCGCCATACCGTTTGTG 540
 Qy 542 TTGTTGGTATGATGACGAATGCTCAAAATTAATCTCATTAATGCTTAGCGCAGTCTC 601
 Db 541 TTGTTGGTATGATGACGAATGCTCAAAATTAATCTCATTAATGCTTAGCGCAGTCTC 600
 Qy 602 TCTATCGCTTCTGAACCCCGGTGCACTGTGCCGAAACGCAAAATGGGGAACACCCGCTT 661
 Db 601 TCTATCGCTTCTGAACCCCGGTGCACTGTGCCGAAACGCAAAATGGGGAACACCCGCTT 660
 Qy 662 TTGGATGATTAATGCAATGCTCCACATTTGATGCTTCCAGATTCGTTGGGAATCTG 721
 Db 661 TTGGATGATTAATGCAATGCTCCACATTTGATGCTTCCAGATTCGTTGGGAATCTG 720
 Qy 722 CTGATAGCTTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTTGACAGCAATA 781
 Db 721 CTGATAGCTTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTTGACAGCAATA 780
 Qy 782 TATAACAGAGGAAGCTGCCCTGTCTTAAACCTTTTATCATCATTTATAGCTTA 841
 Db 781 TATAACAGAGGAAGCTGCCCTGTCTTAAACCTTTTATCATCATTTATAGCTTA 840
 Qy 842 CTTTCATTAATTCGACTGCTTCCAAATGACAAAGCTTTTGATTTTAAAGACATTTTAAAGAC 901
 Db 841 CTTTCATTAATTCGACTGCTTCCAAATGACAAAGCTTTTGATTTTAAAGACATTTTAAAGAC 900
 Qy 902 AACTTTGAGAAGATCAAAAACAACTTAATTTTGAAGGATCCAAACGATGAGATTTCTT 961
 Db 901 AACTTTGAGAAGATCAAAAACAACTTAATTTTGAAGGATCCAAACGATGAGATTTCTT 960
 Qy 962 CAAATTTTACTGCACTTTTATTCGAGCATCTCCGATTAAGCTGCTCCAGTCAACACTA 1021
 Db 961 CAAATTTTACTGCACTTTTATTCGAGCATCTCCGATTAAGCTGCTCCAGTCAACACTA 1020
 Qy 1022 CAACAGAGATGAACGCGCACAAATTCGCGTGAAGCTGTCTATCGGTTTACTCAGATTTAG 1081
 Db 1021 CAACAGAGATGAACGCGCACAAATTCGCGTGAAGCTGTCTATCGGTTTACTCAGATTTAG 1080
 Qy 1082 AAGGGATTTTCGATGTTGCTGTTTCCATTTTCCAGCAGCACAAATAACGGTTATTTGT 1141
 Db 1081 AAGGGATTTTCGATGTTGCTGTTTCCATTTTCCAGCAGCACAAATAACGGTTATTTGT 1140
 Qy 1142 TTATAAATACTACTATTGCGAGCATTTGCTGTAAAGAGAGAGGGGTATCTCTCGAGAAA 1201
 Db 1141 TTATAAATACTACTATTGCGAGCATTTGCTGTAAAGAGAGAGGGGTATCTCTCGAGAAA 1200
 Qy 1202 GAGAGG----- 1207
 Db 1201 GATCGGACCGGCGAGGGGTTCCGGAAGAGGAGGACCCCAAAAAGCTGACCCCTTTAG 1260
 Qy 1208----- 1207
 Db 1261 CTTAAGCAGTTTATCCCAATGTTGGCGGAGAACCCCTAGGCGCCAGCGGAGGTATG 1320
 Qy 1208----- 1207
 Db 1321 AAGGAGATGTTCCAGAAATCCGAGCGATTTAAGGAACTCACCCCAATTAACACCCCG 1380
 Qy 1208----- 1207
 Db 1381 ACATCATATTTAAGGATGAAGAAAACACCGGAGCGGACAGGCTGATGACTCAGAGGTGA 1440
 Qy 1208----- 1207
 Db 1441 AGGACAAGTTGAACGCTTTGGCCATCTCGGTGATGAACAGTGGCGGAGGAGTGAACCTG 1500
 Qy 1208----- 1207

Db 1501 GGGTGACCGAGGGCTGGGACGAAGATGCCACCACTCAGAGAGTCTCTGCACTACGAGG 1560
Qy 1208 ----- 1207
Db 1561 GCGGCGAGTGGACATACACAGCTGACCGGACCGGACGCAAGTACGGCATGTGCGCC 1620
Qy 1208 ----- 1207
Db 1621 GCCTGGCGGTGGAGCGCGCTTCGACTGGGTGTACTAGAGTCCAAAGGCACATATCCACT 1680
Qy 1208 ----- 1207
Db 1681 GCTCGTGAAAGCAGAGAACTCGGTGGCGGCAAAATCGGAGGCGTCGACAAAACCTCACA 1740
Qy 1208 ----- 1207
Db 1741 CATGCCACCGTGGCCACGACCTGAACTCTCTGGGGGACCGTCACTCTTCCTCTTCCGCC 1800
Qy 1208 -----CTGAAGCCAGGCCCGCCAGGCCAGG 1231
Db 1801 CAAAACCAAGACACACCCCTCATGATCTCCCGGACCCCTGAGTCACTGCGTGGTGG 1860
Qy 1232 CCCAGGCCAGGAGAAACATGATCATGCGCCCGCCCGG----- 1269
Db 1861 ACGTGAGCACGAGAGACCCCTGAGGTCAAGTTCAACTGTGTAAGTGGGCGGTGAGGTGC 1920
Qy 1270 -----GAGAGGATAAATTTGGTCTTCCCG 1294
Db 1921 ATAATGCCAAGAACGCGGAGGAGGAGTACACAGACGCTACCGTGTGTCAGCG 1980
Qy 1295 GTGTACCGCCGACGA----- 1310
Db 1981 TCCTCACCGTCTGCAACGAGACTGGCTGTAATGGCAAGGATAAAGTGAAGTCTCCA 2040
Qy 1311 ----- 1310
Db 2041 ACAPAGCCCTCCAGCCCCCATCGAGAAACCACTCTCCAAAGCCAAAGGCGAGCCCGAG 2100
Qy 1311 ----- 1310
Db 2101 AACCACAGGTGTACACCCCTGCCCCCATCCCGGATGAGCTGACCAAGAACAGGTGAGCC 2160
Qy 1311 ---GTGCAGCGAGAGAGTGTCTTTGATGACAGTGTCCGGGATTCCTCGTGTGCT 1366
Db 2161 TGACCTGCTGTGTAAGGCTTCTATCCAGGACATCGCGTGGAGTGGGAGAGCATG 2220
Qy 1367 TCCACCCCATGGCCATCGA----- 1385
Db 2221 GGCAGCCGAGAAACAACTACAGACACGCGCTCGTGTGGAATCCGAGCGCTCCTTCT 2280
Qy 1386 ----- 1385
Db 2281 TCCTCTACAGCAAGTCTACCGTGGCAAGAGAGTGGCAGCGGGAACCTCTTCTCAT 2340
Qy 1386 -----GAACACTCAAGAGAGAAATGTCCCTTTCTAACTAGTGG 1423
Db 2341 GCTCGGTGATGATGAGCTGTGCACACCACTACACGCAAGAGCGCTCTCCCTGTCTC 2400
Qy 1424 CGTAGAATTCCTAGGGCGCGCGGAATTAATTCGCTTAGACATGACTGTTCCTCAGTT 1483
Db 2401 CCGGGAATAGTGGCGGCGCGGAATTAATTCGCTTAGACATGACTGTTCCTCAGTT 2460
Qy 1484 CAAAGTTGGGCACTTACGAGAACCGGCTTCTGCTAGATTCTAATCAAGAGATGTCAGAA 1543
Db 2461 CAAAGTTGGCACTTACGAGAACCGGCTTCTGCTAGATTCTAATCAAGAGATGTCAGAA 2520
Qy 1544 TGCCATTTGCTGAGAGATGCAAGGTTTCAATTTTGGATCTTTTATTTGTAACCTATAT 1603
Db 2521 TGCCATTTGCTGAGAGATGCAAGGTTTCAATTTTGGATCTTTTATTTGTAACCTATAT 2580
Qy 1604 AGTATAGGATTTTCTGCTATTTGTTCTCTGCTAGAGCTTGTCTCTGATCAGCCT 1663

Db 2581 AGTATAGGATTTTCTTGTCAITTTTGTCTTCTGCTACGAGCTTGTCTCTGATCAGCCT 2640
Qy 1664 ATCTCCAGCTGATGAATATCTTGTGTAGGGTTTGGGAAATCAATCGAGTTTGATGT 1723
Db 2641 ATCTCCAGCTGATGAATATCTTGTGTAGGGTTTGGGAAATCAATCGAGTTTGATGT 2700
Qy 1724 TTTTCTGTGATTTCCCACTCTCTCTCAGAGTACAGAAATTAAGTGAAGTTCGTTTG 1783
Db 2701 TTTTCTGTGATTTCCCACTCTCTCTCAGAGTACAGAAATTAAGTGAAGTTCGTTTG 2760
Qy 1784 TGCAAGCTTATCGATGAAGCTTTAATGCGGTAGTTTATCAGGTTAAATTTGCTAAACGACT 1843
Db 2761 TGCAGCTTATCGATGAAGCTTTAATGCGGTAGTTTATCAGGTTAAATTTGCTAAACGACT 2820
Qy 1844 CAGGCAACCGTGTATGAATCTAAACAAATGCGCTCATCTCATCTCGGCAACCGTCAACCTG 1903
Db 2821 CAGGCAACCGTGTATGAATCTAAACAAATGCGCTCATCTCATCTCGGCAACCGTCAACCTG 2880
Qy 1904 GATGCTGTAGGCATAGGCTTGTATGCGGTACTGCGGGGCTCTTGGGGGATATCCTC 1963
Db 2881 GATGCTGTAGGCATAGGCTTGTATGCGGTACTGCGGGGCTCTTGGGGGATATCCTC 2940
Qy 1964 CATTCGACAGCATCGCCAGTCACTATGCGTGTCTCTAGCGCTATATCGTTGATGCAA 2023
Db 2941 CATTCGACAGCATCGCCAGTCACTATGCGTGTCTCTAGCGCTATATCGTTGATGCAA 3000
Qy 2024 TTTTCTATGGCACCCGTTCTCGGAGCACTGTGCGACCGCTTTGGCGCGCCGCGCTCTG 2083
Db 3001 TTTTCTATGGCACCCGTTCTCGGAGCACTGTGCGACCGCTTTGGCGCGCCGCGCTCTG 3060
Qy 2084 CTCGCTTCTGCTACTCTGAGCCACTATCGACTACGCGATCATGCGGACCAACCGCTCTG 2143
Db 3061 CTCGCTTCTGCTACTCTGAGCCACTATCGACTACGCGATCATGCGGACCAACCGCTCTG 3120
Qy 2144 TGGATCTATCGAATCTAAATGTAAGTTAAATCTCTPAAATAATTAATTAAGTCCCAAGTT 2203
Db 3121 TGGATCTATCGAATCTAAATGTAAGTTAAATCTCTPAAATAATTAATTAAGTCCCAAGTT 3180
Qy 2204 CTCATATCAAGAACCTTAACAGCATGCGGTGAGCATCTAGACCTTCAACAGCAGCCAGATC 2263
Db 3181 CTCATATCAAGAACCTTAACAGCATGCGGTGAGCATCTAGACCTTCAACAGCAGCCAGATC 3240
Qy 2264 CATCACTGTCTGGCCAAATATGTTTCTAGTCCCTCAGGAGTTAGCTCTTGTGAAGTGAATA 2323
Db 3241 CATCACTGTCTGGCCAAATATGTTTCTAGTCCCTCAGGAGTTAGCTCTTGTGAAGTGAATA 3300
Qy 2324 CTTCTCGAAGTTGCAAGTGTAACTCCGCTGTATTGACCGGCAATATCCGTTAGTTGGCAA 2383
Db 3301 CTTCTCGAAGTTGCAAGTGTAACTCCGCTGTATTGACCGGCAATATCCGTTAGTTGGCAA 3360
Qy 2384 AGTGTGTTGTTACCGGAGAGTAATCTCCAACTCTCTGAGAGTAGGACCAACAA 2443
Db 3361 AGTGTGTTGTTACCGGAGAGTAATCTCCAACTCTCTGAGAGTAGGACCAACAA 3420
Qy 2444 CACAGATCCAGCGTGTGTACTTGAATCAACATAAGAAAGCAATTCGATTTGCAAGAT 2503
Db 3421 CACAGATCCAGCGTGTGTACTTGAATCAACATAAGAAAGCAATTCGATTTGCAAGAT 3480
Qy 2504 CAAAGTTGTTGAGAGTGTGCAATCACTTTGGGTACAATTTCAACCCCTTGGCAACTGCAAGC 2563
Db 3481 CAAAGTTGTTGAGAGTGTGCAATTTCCAAAGCCTGCTCTGATTTGCAACCCA 3540
Qy 2564 TAGGTTGTAGAGTGTGCAATCACTTTGGGTACAATTTCAACCCCTTGGCAACTGCAAGC 2623
Db 3541 TAGGTTGTAGAGTGTGCAATCACTTTGGGTACAATTTCAACCCCTTGGCAACTGCAAGC 3600
Qy 2624 TTGGTTGTGAACAGCATCTTCAATTTCTGGCAAGCTCTTGTCTGTATATCGACAGCAA 2683
Db 3601 TTGGTTGTGAACAGCATCTTCAATTTCTGGCAAGCTCTTGTCTGTATATCGACAGCAA 3660
Qy 2684 CAGAACTACCTGGGAATCAATCACTGTTGAGCTTGAAGAGTCTGAGGCAACGAA 2743
Db 3661 CAGAACTACCTGGGAATCAATCACTGTTGAGCTTGAAGAGTCTGAGGCAACGAA 3719

QY 2744 ATCTGGATCAGCGTATTTATCAGCAATAAATTAGAACTTTCAGAAAGGCCAGCAGGCATGTC 2803
DB 3720 ATCTGGATCAGCGTATTTATCAGCAATAAATTAGAACTTTCAGAAAGGCCAGCAGGCATGTC 3779
QY 2804 AATACTACAGGGCTGATGTCATTTTGAACCATCATCTTGGCAGCAGTACGAACTG 2863
DB 3780 AATACTACAGGGCTGATGTCATTTTGAACCATCATCTTGGCAGCAGTACGAACTG 3839
QY 2864 GTTTCCTGGACCAATAATTTTGTACACTTAGGAAACAGTTTCTTCCGTAAGCCATAGC 2923
DB 3840 GTTTCCTGGACCAATAATTTTGTACACTTAGGAAACAGTTTCTTCCGTAAGCCATAGC 3899
QY 2924 AGCTACTGCTGGCGGCTCTGCTAGCAGATACATTTAGCAGCAACCTTTGTGGGCAAC 2983
DB 3900 AGCTACTGCTGGCGGCTCTGCTAGCAGATACATTTAGCAGCAACCTTTGTGGGCAAC 3959
QY 2984 GTAGATGACTTCTGGGTAAGGTAAGCATCTTCTTAGGTGGAGTGCATAAAACAATTTTC 3043
DB 3960 GTAGATGACTTCTGGGTAAGGTAAGCATCTTCTTAGGTGGAGTGCATAAAACAATTTTC 4019
QY 3044 TTTGCAACAGCAAACTTTGGCAGGAACAACCCAGCATCAGGGAAGTGAAGGCGCAATTCG 3103
DB 4020 TTTGCAACAGCAAACTTTGGCAGGAACAACCCAGCATCAGGGAAGTGAAGGCGCAATTCG 4079
QY 3104 GGTTCACAGCAATATAGAGGCCAACTTCTCAATAGGTCTTGCRAAAGCAGAGCAGAC 3163
DB 4080 GGTTCACAGCAATATAGAGGCCAACTTCTCAATAGGTCTTGCRAAAGCAGAGCAGAC 4139
QY 3164 TACACCGGCGAAGTCTCAACTTGCAGCGTCTCGGTTAGTTAGCTTTCATGGAAATTTCC 3223
DB 4140 TACACCGGCGAAGTCTCAACTTGCAGCGTCTCGGTTAGTTAGCTTTCATGGAAATTTCC 4199
QY 3224 GACGTATCTATAGAGAGATCAATGGCTCTCTTAAGCTTATCTGGCAATTCGCAATGTC 3283
DB 4200 GACGTATCTATAGAGAGATCAATGGCTCTCTTAAGCTTATCTGGCAATTCGCAATGTC 4259
QY 3284 CTCTGGGAAAGAGCTTCTAAACAGGTGTCCTTCAAGCGACTCCATCAAACTTGGCAGT 3343
DB 4260 CTCTGGGAAAGAGCTTCTAAACAGGTGTCCTTCAAGCGACTCCATCAAACTTGGCAGT 4319
QY 3344 TAGTTCTAAAAGGCTTTGTCAACATTTTGAAGCAATTTGCGCAATTTGGTTGACTAA 3403
DB 4320 TAGTTCTAAAAGGCTTTGTCAACATTTTGAAGCAATTTGCGCAATTTGGTTGACTAA 4379
QY 3404 TTCCATATCTGTCCGTTTCTGTAGAGAGCAGGAGGCATCTTCAATTTCTTGTA 3463
DB 4380 TTCCATATCTGTCCGTTTCTGTAGAGAGCAGGAGGCATCTTCAATTTCTTGTA 4439
QY 3464 GGAGGCTTTAGAAAGCTCAATTTTGCACAAATTAATACGACTTTCAGAAAGGACTTCTTT 3523
DB 4440 GGAGGCTTTAGAAAGCTCAATTTTGCACAAATTAATACGACTTTCAGAAAGGACTTCTTT 4499
QY 3524 AGTTTGGATCTCTTTAGTTGTTCTTGGTATCTGCTGGCTTGGCATCTCCCTTCCCT 3583
DB 4500 AGTTTGGATCTCTTTAGTTGTTCTTGGTATCTGCTGGCTTGGCATCTCCCTTCCCT 4559
QY 3584 TCTAGTACCTTTTAGGACTTCAATATCCAGGTTTCTCTCCACCTCTGCTCCCAACGTCACACC 3643
DB 4560 TCTAGTACCTTTTAGGACTTCAATATCCAGGTTTCTCTCCACCTCTGCTCCCAACGTCACACC 4619
QY 3644 GTACTTGGCAGATCTAACTAATGCAAAATAAATAAGTACGACATATCCCGAGGTATATC 3703
DB 4620 GTACTTGGCAGATCTAACTAATGCAAAATAAATAAGTACGACATATCCCGAGGTATATC 4679
QY 3704 TTCTTTGGATTTAGCTTCTGCAAGTTTCAATAGCTTCTCTCCCTAATTTTAGGCTTCAACA 3762
DB 4680 TTCTTTGGATTTAGCTTCTGCAAGTTTCAATAGCTTCTCTCCCTAATTTTAGGCTTCAACA 4739
QY 3763 AAACCTTCGTCGCAAAATACCGTTTGGTATTAAGAACTTCTTGGAGCAATTTGCTTACGAT 3822
DB 4740 AAACCTTCGTCGCAAAATACCGTTTGGTATTAAGAACTTCTTGGAGCAATTTGCTTACGAT 4799

QY 3823 CCCACAGGTGGCTTCCATGGCTCTAGACCCCTTTGATTTGGCCAAAACAGGAAGTGGCTT 3882
DB 4800 CCCACAGGT-GCTTCCATGGCTCTAGACCCCTTTGATTTGGCCAAAACAGGAAGTGGCTT 4858
QY 3883 CCAAGTGACAGAAACCAACACCTGTTTGTTCACACCAAAATTTCAAGCAGTCTCCATCAC 3942
DB 4859 CCAAGTGACAGAAACCAACACCTGTTTGTTCACACCAAAATTTCAAGCAGTCTCCATCAC 4918
QY 3943 AATCCAAATTCGANTACCCAGAACTTTTTCAGTTGCTCCAGATGTAGCCCTTTATACACA 4002
DB 4919 AATCCAAATTCGANTACCCAGAACTTTTTCAGTTGCTCCAGATGTAGCCCTTTATACACA 4978
QY 4003 AACCTGTACGACGAGATTTGGTAGCTTGTGCTCTTATAGCCCTCCGGAATAGACT 4062
DB 4979 AACCTGTACGACGAGATTTGGTAGCTTGTGCTCTTATAGCCCTCCGGAATAGACT 5038
QY 4063 TTTTGGACGAGTACACAGGCCCAACGAGTAAATTAAGAGTCAAGCCACCAAGTAGTGA 4122
DB 5039 TTTTGGACGAGTACACAGGCCCAACGAGTAAATTAAGAGTCAAGCCACCAAGTAGTGA 5098
QY 4123 ATAGACCAATCGGGCGGTCACTAGTCAAGAGCGCAACAAATTTCACTGACAGGAACT 4182
DB 5099 ATAGACCAATCGGGCGGTCACTAGTCAAGAGCGCAACAAATTTCACTGACAGGAACT 5158
QY 4183 TTTTGACATCTTCAGAAAGTTTCGTATTTCAGTAGTCAATTTGCCGAGCATCAATAATGGGGA 4242
DB 5159 TTTTGACATCTTCAGAAAGTTTCGTATTTCAGTAGTCAATTTGCCGAGCATCAATAATGGGGA 5218
QY 4243 TTATACGAGAGCAACAGTGGAGTCACTCAACACTTTTGGGTCTCAGAAAAGCAT 4302
DB 5219 TTATACGAGAGCAACAGTGGAGTCACTCAACACTTTTGGGTCTCAGAAAAGCAT 5278
QY 4303 AAACAGTTTCTACTACCGCCATTTAGTGAATCTTTCAAAATCGCCAGTGGAGAAAGAAAG 4362
DB 5279 AAACAGTTTCTACTACCGCCATTTAGTGAATCTTTCAAAATCGCCAGTGGAGAAAGAAAG 5338
QY 4363 GCACAGCATATACGANTTAGGGCAAGGATGCAACTTTATCAACGAGGTCTCTATAGA 4422
DB 5339 GCACAGCATATACGANTTAGGGCAAGGATGCAACTTTATCAACGAGGTCTCTATAGA 5398
QY 4423 TAACCTAGCGCTGGGATCATCTTTGCAACAATCTTTCTGCCAAATCTAGTCCAAA 4482
DB 5399 TAACCTAGCGCTGGGATCATCTTTGCAACAATCTTTCTGCCAAATCTAGTCCAAA 5458
QY 4483 TCACCTATTGATACCATTTATGTAACAATTTGAGCAAGTTGTGATCAGTCTCTCAAAAT 4542
DB 5459 TCACCTATTGATACCATTTATGTAACAATTTGAGCAAGTTGTGATCAGTCTCTCAAAAT 5518
QY 4543 GGTCTCTGTAAAGGATGACTCAACTTCACATTAACCTTGAAGTCACTCAATTTGAGTGA 4602
DB 5519 GGTCTCTGTAAAGGATGACTCAACTTCACATTAACCTTGAAGTCACTCAATTTGAGTGA 5578
QY 4603 ACTTGATCAGGTTGTGACGCTGGTCAAGCAGCATAGGAAACAACGCTTTTCTACCAAA 4662
DB 5579 ACTTGATCAGGTTGTGACGCTGGTCAAGCAGCATAGGAAACAACGCTTTTCTACCAAA 5638
QY 4663 TCAGGAATATCAAACTCTGCAACACTTGGTATGAGTGAAGGAAATGTCATAC 4722
DB 5639 TCAGGAATATCAAACTCTGCAACACTTGGTATGAGTGAAGGAAATGTCATAC 5698
QY 4723 TTGAAGTCGACAGTGAAGTGTAGTCTTGAGAAATTTCTGAAGCCGTTATTTATTCAGT 4782
DB 5699 TTGAAGTCGACAGTGAAGTGTAGTCTTGAGAAATTTCTGAAGCCGTTATTTATTCAGT 5758
QY 4783 GAGTCAGTCATCAGGAGATCTCTACCGCGGAGGATCTGTGGCCGACCTGACGTCGGA 4842
DB 5759 GAGTCAGTCATCAGGAGATCTCTACCGCGGAGGATCTGTGGCCGACCTGACGTCGGA 5818
QY 4843 TCACCGGCGCCACAGGTGGGTGCTGCGGCTTATATCGCGGACATCAACCGATGGGGAAG 4902
DB 5819 TCACCGGCGCCACAGGTGGGTGCTGCGGCTTATATCGCGGACATCAACCGATGGGGAAG 5878
QY 4903 ATCGGCTCGCCACCTTCGGGCTCATAGCGCTTTGTTTGGCGGTGGGTATGTTGGCAGGCC 4962

Db	5879		ATCGGGCTCGCCACTTCGGGCTCATGACGCGTTGTTTCGGCGTGGGTATGGTGGCAGGCC	5938
Qy	4963		CCGTGGCCGGGGGACTGTTGGGGCCCATCTCTTTGCAATGCAC	5004
Db	5939		CCGTGGCCGGGGGACTGTTGGGGCCCATCTCTTTGGACCTCGACGGGGGGGGGNAAG	5998
Qy	5005	-----	-----	5004
Db	5999		CCAGTTGTGTCTCAAAATCTCTGATGTTACATTTGCACAAAGATAAAATATATCATCATG	6058
Qy	5005	-----	-----	5004
Db	6059		AACAATAAACTGCTCTGCTTACATAAACAAGTAACAGTAAGGGGTGTTATGAGCCCATATTC	6118
Qy	5005	-----	-----	5004
Db	6119		ACGGGAACGTCCTTCTCAAGCCGCGATTAAATTCAAACATGATGCTGATTATATATGG	6178
Qy	5005	-----	-----	5004
Db	6179		GTATAAATGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAATCTATCGAATTTGATGG	6238
Qy	5005	-----	-----	5004
Db	6239		GAAGCCGATGGCCAGAGTTGTTCTGAAACATGGCAAGGTAGCGTGGCCATGATGT	6298
Qy	5005	-----	-----	5004
Db	6299		TACAGATGAGATGGTTCAGACTAAACTGGCTGACGGAAATTTATGCTCTTCGACCATCAA	6358
Qy	5005	-----	-----	5004
Db	6359		GCAATTTATCCGTACTCTCGATGATGATGATGATGATGATGATGATGATGATGATGATG	6418
Qy	5005	-----	-----	5004
Db	6419		AGCAATTCAGATTTATGAGAATAATCTGATTCAGGTGAAAATATTTGTTGATGCGCTGGC	6478
Qy	5005	-----	-----	5004
Db	6479		AGTGTCTCGCCGGTTCATTCGATTCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	6538
Qy	5005	-----	-----	5004
Db	6539		CGTATTTCTGCTCGCTCAGGCGCATCACGAATGAATACCGTTTGGTTGATGCGAGTGA	6598
Qy	5005	-----	-----	5004
Db	6599		TTTTGATGACGAGCGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	6658
Qy	5005	-----	-----	5004
Db	6659		TTTGCCATTTCCACCGGATTCAGTCGTCATCTGATGATGATGATGATGATGATGATGATG	6718
Qy	5005	-----	-----	5004
Db	6719		TTTTGACGAGGGAAATTAATAGTTGTTATGATGTTGAGTGGAGTCGGAATCGCAGACCG	6778
Qy	5005	-----	-----	5004
Db	6779		ATACAGGATCTTGCCATCTCTATGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	6838
Qy	5005	-----	-----	5004
Db	6839		ACGGCTTTTCAAAAATATGGTATTCGATAATCTCTGATGATAAATTCAGTTTCAATTT	6898
Qy	5005	-----	-----	5004
Db	6899		GATGCTCGATGAGTTTTTCTAATCAGAAATGGTTAAATGGTTGTAACACTGCGCAGACAT	6958
Qy	5005	-----	-----	5004

Db	6959		TACGCTGACTTGACGGGACGGCGGCTTGTGTAATAAATCGAACTTTTGCTGAGTTGAAG	7018
Qy	5005	-----	-----	5004
Db	7019		GATCAGATCAGCATCTTCCGACAAACGAGACCGTTCCGTGGCAAGCAAAAGTTCAA	7078
Qy	5005	-----	-----	5004
Db	7079		ATCACAACTGGTGCACCTACAAAGAGCTCTCATCAACCGTGGCTCCCTCACTTCTCGG	7138
Qy	5005	-----	-----	5004
Db	7139		CTGGATGATGGGGCGATTTCAGGCTGTATGATGATGATGATGATGATGATGATGATGATG	7198
Qy	5005	-----	-----	5004
Db	7199		CTCAGCGCCCCCCCCCTCAGGTCACCGGGGGGGTCTCAACGGCTCAACCTA	7258
Qy	5043		CTACTGGGCTGCTTCTTAATGAGGAGTCCGATAGGAGGAGCGTCGAGTATCTATGATT	5102
Db	7259		CTACTGGGCTGCTTCTTAATGAGGAGTCCGATAGGAGGAGCGTCGAGTATCTATGATT	7318
Qy	5103		GGAGTATGGGAATGATGATACCGGATTTCTCAGTGTCTTGAAGTCTCTCATCAGATT	5162
Db	7319		GGAGTATGGGAATGATGATACCGGATTTCTCAGTGTCTTGAAGTCTCTCATCAGATT	7378
Qy	5163		TGCCCAACTAAAGCAACCGGAGGAGGATTTCTCATGTAATTTCTCATCTTTTGGTCA	5222
Db	7379		TGCCCAACTAAAGCAACCGGAGGAGGATTTCTCATGTAATTTCTCATCTTTTGGTCA	7438
Qy	5223		TCAGTATGACTGCAACTGTGAGACTATCTCGTTATGACAGCAGAAATGCTCTTTGGAG	5282
Db	7439		TCAGTATGACTGCAACTGTGAGACTATCTCGTTATGACAGCAGAAATGCTCTTTGGAG	7498
Qy	5283		ACAGTAAATGAAAGTCCCAACCAATGAGTGGTATGAGTGGTATGAGTGGTATGAGTGG	5342
Db	7499		ACAGTAAATGAAAGTCCCAACCAATGAGTGGTATGAGTGGTATGAGTGGTATGAGTGG	7558
Qy	5343		CGAACTTTTTCGGTGCCTTGAACTATAAAATGAGTGGTATGAGTGGTATGAGTGGTATG	5402
Db	7559		CGAACTTTTTCGGTGCCTTGAACTATAAAATGAGTGGTATGAGTGGTATGAGTGGTATG	7618
Qy	5403		CGGGGCAAAATGCTTACCTCTGACCTTCAAGAGGATGATGAGGTTTGTAGTACTGATG	5462
Db	7619		CGGGGCAAAATGCTTACCTCTGACCTTCAAGAGGATGATGAGGTTTGTAGTACTGATG	7678
Qy	5463		CCAACTTCAGTGACAAAGTGTCTTATTCGTTCAAAACCATTCGAAATCCAGAGAAATCAA	5522
Db	7679		CCAACTTCAGTGACAAAGTGTCTTATTCGTTCAAAACCATTCGAAATCCAGAGAAATCAA	7738
Qy	5523		GTTTGTGTCTACTATTGATCCAAAGCCAGTGGGCTTGAACCTGACAATGATGCTCG	5582
Db	7739		GTTTGTGTCTACTATTGATCCAAAGCCAGTGGGCTTGAACCTGACAATGATGCTCG	7798
Qy	5583		TGTTTGTAGGTCTCTTGTATGAAATGATGATGATGATGATGATGATGATGATGATGATG	5642
Db	7799		TGTTTGTAGGTCTCTTGTATGAAATGATGATGATGATGATGATGATGATGATGATGATG	7858
Qy	5643		CCAAAGCGATAAATACCAAACTCTTTTAAACCTCTTTTAAACCTCTTTTAAAGCAAGTATG	5702
Db	7859		CCAAAGCGATAAATACCAAACTCTTTTAAACCTCTTTTAAACCTCTTTTAAAGCAAGTATG	7918
Qy	5703		TGCTGTATTAACCCCAATCAGCTCGTAGTCTGATCTCTCATCACTTTGAGGGGCACTA	5762
Db	7919		TGCTGTATTAACCCCAATCAGCTCGTAGTCTGATCTCTCATCACTTTGAGGGGCACTA	7978
Qy	5763		CTTTGTTTATAGAAATTTTCGGGAGATGCGATATCGAGAAAAAGGTACGCTGATTTTAA	5822
Db	7979		CTTTGTTTATAGAAATTTTCGGGAGATGCGATATCGAGAAAAAGGTACGCTGATTTTAA	8038
Qy	5823		CGTGAATTTTATCTCAAGATCTCTGCTCGCGGTTTGGTGTATGATGCGTGAACCTCT	5882
Db	8039		CGTGAATTTTATCTCAAGATCTCTGCTCGCGGTTTGGTGTATGATGCGTGAACCTCT	8098

Db 10259 TCCCGAAAGTGCACCTGACGCTAAGAAACCAATATTATCATGACATTAACCTATAA 10318
 Qy 8103 AAATAGGCGTATACAGAGCCCTTTTCGCTCTTCAAGAAATTAATCTCATGTTTGACAGCTT 8162
 Db 10319 AAATAGGCGTATACAGAGCCCTTTTCGCTCTTCAAGAAATTAATCTCATGTTTGACAGCTT 10378
 Qy 8163 ATCATCGATAGCTGACATCATGTTGGTATTGTGAATAGAGCGAGATCGGGAACACTGAA 8222
 Db 10379 ATCATCGATAGCTGACATCATGTTGGTATTGTGAATAGAGCGAGATCGGGAACACTGAA 10438
 Qy 8223 AAATAACAGTTATTATTTCG 8241
 Db 10439 AAATAACAGTTATTATTTCG 10457
 RESULT 14
 LOCUS AX146619 10492 bp DNA linear PAT 31-MAY-2001
 DEFINITION Sequence 81 from Patent WO0134654.
 ACCESSION AX146619
 VERSION AX146619.1 GI:14295012
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Strauch, K.
 TITLE Hedgehog fusion proteins and uses
 JOURNAL Patent: WO 0134654-A 81 17-MAY-2001;
 BIOGEN, INC. (US)
 FEATURES
 source 1..10492
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 68.9%; Score 5677.8; DB 6; Length 10492;
 Best Local Similarity 77.2%; Pred. No. 0;
 Matches 8101; Conservative 0; Mismatches 137; Indels 2251; Gaps 9;
 Qy 2 GATCTAACATCCAAAGAGAAAGGTGAAATGAATGAACCTTTTGCCATCCGACATCCACAGG 61
 Db 1 GATCTAACATCCAAAGAGAAAGGTGAAATGAATGAACCTTTTGCCATCCGACATCCACAGG 60
 Qy 62 TCCATCTCACATTAAGTCCAAAGCGACAGAGGGGATACACTAGCAGAGACCGCTT 121
 Db 61 TCCATCTCACATTAAGTCCAAAGCGACAGAGGGGATACACTAGCAGAGACCGCTT 120
 Qy 122 GCAAGCGAGGACCTCCACTCTCTCTCTCAACACCCACCTTTTGCCATCGAAAAACCA 181
 Db 121 GCAAGCGAGGACCTCCACTCTCTCTCAACACCCACCTTTTGCCATCGAAAAACCA 180
 Qy 182 GCCAGATTATGGGCTTGATTTGAGCTCGCTCATTCCTCAATTCCTTCTATTAGGCTACTAA 241
 Db 181 GCCAGATTATGGGCTTGATTTGAGCTCGCTCATTCCTCAATTCCTTCTATTAGGCTACTAA 240
 Qy 242 CACCATGACTTTATTAGCTGCTATCTCTGGCCCTCGCGAGGTTCAATGTTTGTAT 301
 Db 241 CACCATGACTTTATTAGCTGCTATCTCTGGCCCTCGCGAGGTTCAATGTTTGTAT 300
 Qy 302 TTCGAATGCAACAAAGCTCGCATTTACCCGAAACATCTCCAGATGAGGGCTTTCTGA 361
 Db 301 TTCGAATGCAACAAAGCTCGCATTTACCCGAAACATCTCCAGATGAGGGCTTTCTGA 360
 Qy 362 GTGTGGGTCAATAGTTTCATGTTCCCAATGGCCCAAACTGACAGTTTAAAGCTG 421
 Db 361 GTGTGGGTCAATAGTTTCATGTTCCCAATGGCCCAAACTGACAGTTTAAAGCTG 420
 Qy 422 TCTTGGAACTTAATATGACAAAAGCGTGATCTCATCCAAAGATGAACATTAAGTTTGGTTCTG 481

Db 421 TCTTGGAACTTAATATGACAAAAGCGTGATCTCATCCAAGATGAACATAAGTTTGGTTCTG 480
 Qy 482 TGAATGCTAAAGCGCCAGTTGTCATAAAGAACTTCGAAAGATCGCCATACCGTTTGC 541
 Db 481 TGAATGCTAAAGCGCCAGTTGTCATAAAGAACTTCGAAAGATCGCCATACCGTTTGC 540
 Qy 542 TTGTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAATGCTTAGCGCAGTCTC 601
 Db 541 TTGTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAATGCTTAGCGCAGTCTC 600
 Qy 602 TCTATCGCTTCTGAACCCCGGTGCACCTGTCGGAACGCAATGGGGAACACCGCTT 661
 Db 601 TCTATCGCTTCTGAACCCCGGTGCACCTGTCGGAACGCAATGGGGAACACCGCTT 660
 Qy 662 TTTGGATGATTATGCAATGCTCCCATTTGATGCTTCCAGATTTCTGTTGGGAATACG 721
 Db 661 TTTGGATGATTATGCAATGCTCCCATTTGATGCTTCCAGATTTCTGTTGGGAATACG 720
 Qy 722 CTGATAGCCTAAGTTTCATGATCAAAATTTAATCTTAAACCTTCTTAAACCTTCTTAAACCTT 781
 Db 721 CTGATAGCCTAAGTTTCATGATCAAAATTTAATCTTAAACCTTCTTAAACCTTCTTAAACCTT 780
 Qy 782 TATAACAGAGGAAGCTGCCCTGCTTAAACCTTCTTAAACCTTCTTAAACCTTCTTAAACCTT 841
 Db 781 TATAACAGAGGAAGCTGCCCTGCTTAAACCTTCTTAAACCTTCTTAAACCTTCTTAAACCTT 840
 Qy 842 CTTTTCATATTGCGACTGTTCCAAATTGAACAAGCTTTTGAATTTAAAGCTTTTAAAGCTT 901
 Db 841 CTTTTCATATTGCGACTGTTCCAAATTGAACAAGCTTTTGAATTTAAAGCTTTTAAAGCTT 900
 Qy 902 AACTTGAGAAGATCAAAAACAACTAATTTTCGAGAGATCCAAACGATGAGATTTCCCTT 961
 Db 901 AACTTGAGAAGATCAAAAACAACTAATTTTCGAGAGATCCAAACGATGAGATTTCCCTT 960
 Qy 962 CAATTTTACTGCGACTTTTATTCGAGCATCTCCGCTAGCTGCTCCAGTCAACACTA 1021
 Db 961 CAATTTTACTGCGACTTTTATTCGAGCATCTCCGCTAGCTGCTCCAGTCAACACTA 1020
 Qy 1022 CAACAGAGATGAAGCAACAAATTCGCGCTGAAGCTGATCGTCTGATCGATTTAG 1081
 Db 1021 CAACAGAGATGAAGCAACAAATTCGCGCTGAAGCTGATCGTCTGATCGATTTAG 1080
 Qy 1082 AAGGGATTTGATGTTGCTGTTTTCGCTTTTCCAAACAGCAAAATAACGGGTATTGT 1141
 Db 1081 AAGGGATTTGATGTTGCTGTTTTCGCTTTTCCAAACAGCAAAATAACGGGTATTGT 1140
 Qy 1142 TTATAAATACTACTATTTCGAGCATTTCTCTAAAGAGAGGGGTATCTCTCGAGAAA 1201
 Db 1141 TTATAAATACTACTATTTCGAGCATTTCTCTAAAGAGAGGGGTATCTCTCGAGAAA 1200
 Qy 1202 GAGGGCTGAAGCCCGG----- 1219
 Db 1201 GATCGGACCGGCGAGGGGTTGCGGAAGAGAGGAGGACCCCAAAAGCTGACCCCTTTAG 1260
 Qy 1220 -----CCGAGGCCAGGCCCGCCAGGCCAG 1242
 Db 1261 CCTACAGCAGTTTATCCCAATGTTGCGCGAGAGAACCTTAGCGCCAGCGGAGGTATG 1320
 Qy 1243 GAAGAAACAT----- 1252
 Db 1321 AAGGGAAGATCTCCGAAACTCCGAGCGATTTAAGGAATCAACCCCAATTAACCCCG 1380
 Qy 1253 -----GTATCATGGCCCCCGGAGAGGATAAATTTGTGCTTCCCCGGGTGTC 1299
 Db 1381 ACATCATATTAAAGATTGAAGAAACACCGGAGCGGAGCTGATGACTCAGAGGTGTA 1440
 Qy 1300 A----- 1300
 Db 1441 AGGCAAGTTGAAGCTTTGGCCATCTCGGTGATGAACACAGTGGCCAGGAGTGAACCTGC 1500
 Qy 1301 ----- 1300
 Db 1501 GGGTGACCGAGGCTGGAGCGAAGATGGCCACCACTCAGAGGAGTCTCTGCACCTACGAGG 1560

QY 1301 ----- 1300
Db 1561 GCCGCGAGTGGACATACACAGTCTGTGACCGCGACCGCAGCAAGTACGGCATGTGCGCC 1620
QY 1301 ----- 1300
Db 1621 GCCTGGCGGTGGAGCGCGCTTCGACTGGGTGTACTAGAGTCCAAGGCACATATCCACT 1680
QY 1301 ----- 1300
Db 1681 GCTCGGTGAAGCAGAGAACTCGGTGGCGGCGAAATCGGAGGCGTGCAGTGGCCAGG 1740
QY 1301 ----- 1300
Db 1741 ATTGTGGTTGAAGCCTTGCAATGTACAGTCCCAGAAAGTATCATCTGTCTCATCTCC 1800
QY 1301 ----- 1300
Db 1801 CCCAAAGCCCAAGGATGTCTCACATTACTCTGACTCTCTTAAGTCACTGTGTGTGG 1860
QY 1301 ----- 1300
Db 1861 TAGACATCAGCAAGGATGATCCGAGGTCCAGTTCAGTGTGTTGTAGATGTGGAGG 1920
QY 1301 ----- 1300
Db 1921 TGCAACAGCTCAGACGCAACACGCGGAGAGAGAGTTCCTCAAGCACATTCGCTCAGTCA 1980
QY 1301 ----- 1300
Db 1981 GTGAATCTCCCATCATGACACGAGAGTGGTCTCAATGGCAAGGAGTTCAAAATGCAAGGTCA 2040
QY 1301 ----- 1300
Db 2041 ACAGTGCAGCTTCCTCGCCCCATCGAGAAACCAATCTCCAAACCAAGGACAGCGCA 2100
QY 1301 ----- CCG 1303
Db 2101 AGGCTCCACAGTGTACACATTCACCTCCCAAGGAGCAGATGGCCAAAGGATAAAGTCA 2160
QY 1304 CCCACAGTGCAGGAGAGAGTGTCTGTTTGTATGACAGTGCAGGAGATTCCTCGGT 1363
Db 2161 GTCTGACCTGATGATATACAGATCTCTCCCTGAAGACATTTACTGTGGAGTGGCAGTGA 2220
QY 1364 GCTCCACCCCATGCGCCATCGAGAAC----- 1390
Db 2221 ATGGGACGCGGAGAGAACTACAAGAACACTCAGCCCATCATGGACACAGATGGCTCTT 2280
QY 1391 ----- 1390
Db 2281 ACTTCGTCTACAGCAAGCTCAATGTGCAAGAGCAACTGGGAGGCGAGAAATCTTTCA 2340
QY 1391 ----- CTCAAGAGAGAGATGTCCCTTTCTAACT 1418
Db 2341 CCTGCTCTGTGTATCAGGGCTGTCACACACCACCTACTGAGAGAGCTCTCCCACT 2400
QY 1419 AGTGGCGTGA-----ATTCCCTAGGCGCGCGGAATTA 1453
Db 2401 CTCCTGTTAAATGATCCAGTGTCTTGGAGCCCTCTGTGCTCTACAGCGCGCGGAATTA 2460
QY 1454 ATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTGGGCACTTACAGAGAGACCGGTCT 1513
Db 2461 ATTCGCTTAGACATGACTGTTCTCAGTTCAAGTTGGGCACTTACAGAGAGACCGGTCT 2520
QY 1514 TGCTAGATCTTAATCAAGAGAGATGTCAAGTGCATTTGCTGAGAGATGCGAGCTTCAT 1573
Db 2521 TGCTAGATCTTAATCAAGAGAGATGTCAAGTGCATTTGCTGAGAGATGCGAGCTTCAT 2580
QY 1574 TTTTGATACCTTTTATTTGTAACCTATATAGTATAGATTTTGTCTATTTTGTTC 1633
Db 2581 TTTTGATACCTTTTATTTGTAACCTATATAGTATAGATTTTGTCTATTTTGTTC 2640

QY 1634 TTCTGTAGAGCTTGCTCTGATCAGCCTATCTCGAGCTGATGAATATCTTGTGGTAG 1693
Db 2641 TTCTGTAGAGCTTGCTCTGATCAGCCTATCTCGAGCTGATGAATATCTTGTGGTAG 2700
QY 1694 GGGTTTGGAAATCATTTGAGTTTGTGTTTCTTGGTATTTCCCACTCCTCTTTCAGA 1753
Db 2701 GGGTTTGGAAATCATTTGAGTTTGTGTTTCTTGGTATTTCCCACTCCTCTTTCAGA 2760
QY 1754 GTACAGAGATTAAGTGAAGTTCGTTTGTGCAAGCTTATCGATAAGCTTTAATGCGGT 1813
Db 2761 GTACAGAGATTAAGTGAAGTTCGTTTGTGCAAGCTTATCGATAAGCTTTAATGCGGT 2820
QY 1814 AGTTTATCAAGTTAAATTTGCTTAACGCGAGTCAGGACCGTGTATGAAATCTAACAATCG 1873
Db 2821 AGTTTATCAAGTTAAATTTGCTTAACGCGAGTCAGGACCGTGTATGAAATCTAACAATCG 2880
QY 1874 CTCATCGTCTCTCGGACCGTCACTCGTGTGATGCTGTAGCATAGGCTTGGTTATGCGG 1933
Db 2881 CTCATCGTCTCTCGGACCGTCACTCGTGTGATGCTGTAGCATAGGCTTGGTTATGCGG 2940
QY 1934 GTACTGCCGGGCTCTTGGGATATCGTCCATTCGACAGCATCGCCAGTCACTATGCG 1993
Db 2941 GTACTGCCGGGCTCTTGGGATATCGTCCATTCGACAGCATCGCCAGTCACTATGCG 3000
QY 1994 GTGCTGTAGCTATATGCGTTGATGCAATTTCTATGCGCACCGCTTCTCGAGCACTG 2053
Db 3001 GTGCTGTAGCTATATGCGTTGATGCAATTTCTATGCGCACCGCTTCTCGAGCACTG 3060
QY 2054 TCGGACCGTTTGGCGCGCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2113
Db 3061 TCGGACCGTTTGGCGCGCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3120
QY 2114 TAGCGATCATGGCAGCACACCGCTCTGTGATGCTATCGAATCTAAATGTAGTTAAA 2173
Db 3121 TAGCGATCATGGCAGCACACCGCTCTGTGATGCTATCGAATCTAAATGTAGTTAAA 3180
QY 2174 ATCTCTAATTAATTAATTAAGTCCAGTTTCTCCATACGAACTTAAACAGATGCGGTG 2233
Db 3181 ATCTCTAATTAATTAATTAAGTCCAGTTTCTCCATACGAACTTAAACAGATGCGGTG 3240
QY 2234 AGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAAATATGTTTCACTCC 2293
Db 3241 AGCATCTAGACCTTCAACAGCAGCAGATCCATCACTGCTTGGCCAAATATGTTTCACTCC 3300
QY 2294 CTCAGAGTTAGCTCTGTGAGTGAATGAACTTCTGGAAGTTCAGTGTAACTCCGCT 2353
Db 3301 CTCAGAGTTAGCTCTGTGAGTGAATGAACTTCTGGAAGTTCAGTGTAACTCCGCT 3360
QY 2354 GTATTGACGGGATATCCGTACGTTGGCAAGTGTGTTGTTACCGAGGAGTAACTCTCC 2413
Db 3361 GTATTGACGGGATATCCGTACGTTGGCAAGTGTGTTGTTACCGAGGAGTAACTCTCC 3420
QY 2414 ACACTCTCTGAGAGTAGGACCAACAAACACAGATCCAGGCTTGTACTGTATCAAC 2473
Db 3421 ACACTCTCTGAGAGTAGGACCAACAAACACAGATCCAGGCTTGTACTGTATCAAC 3480
QY 2474 ATAAGAAGAGCATTCCTGATTGAGGATCAAGTGTTCAGGAGCTACTGATTGGACAT 2533
Db 3481 ATAAGAAGAGCATTCCTGATTGAGGATCAAGTGTTCAGGAGCTACTGATTGGACAT 3540
QY 2534 TTCCAAAGCCTCTCGTAGTTGCAACCGATAGGTTGTAGAGTGTGCAATCACTTCCG 2593
Db 3541 TTCCAAAGCCTCTCGTAGTTGCAACCGATAGGTTGTAGAGTGTGCAATCACTTCCG 3600
QY 2594 TACAATTTCAACCTTGGCACTGACAGCTTGGTTGTGGAACAGCATCTTCAATCTGCG 2653
Db 3601 TACAATTTCAACCTTGGCACTGACAGCTTGGTTGTGGAACAGCATCTTCAATCTGCG 3660
QY 2654 AAGCTCCTTGTCTGATATGACAGGCAACAGAAATCACTGGGAATCAATACCATGTTTC 2713
Db 3661 AAGCTCCTTGTCTGATATGACAGGCAACAGAAATCACTGGGAATCAATACCATGTTTC 3720
QY 2714 AGCTTGACAGAGAGGTCTGAGGCAACGAAATCTGGATCAGCGTATTTATCAGCAATAAC 2773

Db	3721	AGCTTGAG-CAGAGGCTGAGGCAACGAATCTGGATCAGGCTATTATACAGCAATAAC	3779	Db	4800	AAGAACTCTTCGAGCATCTCTTACGATCCCAAGGT-GCTTCCATGGCTTAAGAC	4858
Qy	2774	TAGAACTTCAGAGGCCAGCAGGATGTCATTAATCTACACAGGCTGATGTGCTATTG	2833	Db	3853	CTTTGATTGGCCAAAACAGGAAGTCGTTTCAAGTGCAGAGAAACCAACACTGTTTGT	3912
Db	3780	TAGAACTTCAGAGGCCAGCAGGATGTCATTAATCTACACAGGCTGATGTGCTATTG	3839	Qy	4859	CTTTGATTGGCCAAAACAGGAAGTCGTTTCAAGTGCAGAGAAACCAACACTGTTTGT	4918
Qy	2834	AACCATCATCTTGGCAGCAGTAACGAACCTGTTTCTTGGACCAATAATTTTGTACACTT	2893	Db	3913	CAACCAAAATTTCAAGCAGTCTCATCAATCCAAATTCGATACCCAGCAACTTTTGA	3972
Db	3840	AACCATCATCTTGGCAGCAGTAACGAACCTGTTTCTTGGACCAATAATTTTGTACACTT	3899	Db	4919	CAACCAAAATTTCAAGCAGTCTCATCAATCCAAATTCGATACCCAGCAACTTTTGA	4978
Qy	2894	AGGAACAGTTTCTGTTCCGTAAGCAATAGCAGTACTGCTGGGGGCTCTGCTAGCAC	2953	Qy	3973	TTGCTCCAGATGATGACCTTTATACCAAAACCGTGACGAGAGATTGGTAGACTCCAG	4032
Db	3900	AGGAACAGTTTCTGTTCCGTAAGCAATAGCAGTACTGCTGGGGGCTCTGCTAGCAC	3959	Db	4979	TTGCTCCAGATGATGACCTTTATACCAAAACCGTGACGAGAGATTGGTAGACTCCAG	5038
Qy	2954	GATACACTTAGCAACCACTTGTGGCAACCTGATGACTTCTGGGGTAAGGGTACCATC	3013	Qy	4033	TTTGTCTCTTATAGCCTCCGGAATAGACTTTTGGACGAGTACACCGGCCCAACGAGT	4092
Db	3960	GATACACTTAGCAACCACTTGTGGCAACCTGATGACTTCTGGGGTAAGGGTACCATC	4019	Db	5039	TTTGTCTCTTATAGCCTCCGGAATAGACTTTTGGACGAGTACACCGGCCCAACGAGT	5098
Qy	3014	CTTCTTAGTGTGAGATGCAAAAACAAATTTCTTTCGAAACGACAACTTTGGCAGGAACAC	3073	Qy	4093	AATTAGAGAGTCAGCAGCAGCAACAAAGTAGTGAATAGACCATCGGGCGGTCAAGTCAAAAG	4152
Db	4020	CTTCTTAGTGTGAGATGCAAAAACAAATTTCTTTCGAAACGACAACTTTGGCAGGAACAC	4079	Db	5099	AATTAGAGAGTCAGCAGCAGCAACAAAGTAGTGAATAGACCATCGGGCGGTCAAGTCAAAAG	5158
Qy	3074	CAGCATAGGGAAGTGGAGGAGCAATTTGGGTTCCACAGCAATATAGAGGCCAATTT	3133	Qy	4153	ACGCCAACAAAATTTCACTGACAGGGGAACCTTTTGTGACATCTTCAGAAAAGTTCGTATTCAG	4212
Db	4080	CAGCATAGGGAAGTGGAGGAGCAATTTGGGTTCCACAGCAATATAGAGGCCAATTT	4139	Db	5159	ACGCCAACAAAATTTCACTGACAGGGGAACCTTTTGTGACATCTTCAGAAAAGTTCGTATTCAG	5218
Qy	3134	CTCAATAGTCTTGCAAAACGAGAGCAGCTACACAGGCAAGTCTCAACTTGCACAGT	3193	Qy	4213	TAGTCAATTTGCCAGCATCAATTAATGGGATTTATCCAGAGCAACAGTGGAGTCAACAT	4272
Db	4140	CTCAATAGTCTTGCAAAACGAGAGCAGCTACACAGGCAAGTCTCAACTTGCACAGT	4199	Db	5219	TAGTCAATTTGCCAGCATCAATTAATGGGATTTATCCAGAGCAACAGTGGAGTCAACAT	5278
Qy	3194	CTCCGTTAGTTGAGTTTCAGAAATTTCTGACGTTATCTATAGAGATCAATGGCTCT	3253	Qy	4273	CTACCAACTTTTGGGCTCTCAGAAAAAGCATAAACAGTCTTACTACCGCCATTAAGTGAAC	4332
Db	4200	CTCCGTTAGTTGAGTTTCAGAAATTTCTGACGTTATCTATAGAGATCAATGGCTCT	4259	Db	5279	CTACCAACTTTTGGGCTCTCAGAAAAAGCATAAACAGTCTTACTACCGCCATTAAGTGAAC	5338
Qy	3254	CTTACAGTTATCTGCAATGTCATTAAGTTCTTCTGGGAAAGAGCTTCTAAACAGGTTG	3313	Qy	4333	TTTTCAAATCGCCAGTGGAGAGAAAAGGCAACAGGATCTAGCATTAGCGGCAAGG	4392
Db	4260	CTTACAGTTATCTGCAATGTCATTAAGTTCTTCTGGGAAAGAGCTTCTTACACAGGTTG	4319	Db	5339	TTTTCAAATCGCCAGTGGAGAGAAAAGGCAACAGGATCTAGCATTAGCGGCAAGG	5398
Qy	3314	CTTCAAGAGCACTCCATCAAACTTGGCAGTTAGTTCTTAAAGGGCTTTGTCAACATTTG	3373	Qy	4393	ATGCAACTTTTATCAACAGGGTCTCTATAGATAACCTTAGCCCTGGATCATCTTTTGA	4452
Db	4320	CTTCAAGAGCACTCCATCAAACTTGGCAGTTAGTTCTTAAAGGGCTTTGTCAACATTTG	4379	Db	5399	ATGCAACTTTTATCAACAGGGTCTCTATAGATAACCTTAGCCCTGGATCATCTTTTGA	5458
Qy	3374	AGCAATTTGTCAGCAATTTGGTTTGCATTAATTCATATCTGTTCCGTTTCTGGATAGG	3433	Qy	4453	CAACTCTTTCTGCAAAATCTAGGTCGAAATCTCTTATTTGATACCATTTTGTACAACT	4512
Db	4380	AGCAATTTGTCAGCAATTTGGTTTGCATTAATTCATATCTGTTCCGTTTCTGGATAGG	4439	Db	5459	CAACTCTTTCTGCAAAATCTAGGTCGAAATCTCTTATTTGATACCATTTTGTACAACT	5518
Qy	3434	ACGACGAGGGATCTTCAATTTCTGTGAGGAGGCTTGAAGAGCTCAATTTTGCACAA	3493	Qy	4513	TGACCAAGTTTGTGATCAGCTCTCTCAATTTGGTCTCTGTAACGGATGATCAACTTGCA	4572
Db	4440	ACGACGAGGGATCTTCAATTTCTGTGAGGAGGCTTGAAGAGCTCAATTTTGCACAA	4499	Db	5519	TGACCAAGTTTGTGATCAGCTCTCTCAATTTGGTCTCTGTAACGGATGATCAACTTGCA	5578
Qy	3494	TTCAATAGCACTTCAGAGGAGCTTCTTTAGGTTTGGATTCCTTTAGGTTTGTCTT	3553	Qy	4573	CATTAACTTGAAGCTCAGTCTGATTTGATGAACTTTGATCAGGTTTGTGAGTGTGTCAGAG	4632
Db	4500	TTCAATAGCACTTCAGAGGAGCTTCTTTAGGTTTGGATTCCTTTAGGTTTGTCTT	4559	Db	5579	CATTAACTTGAAGCTCAGTCTGATTTGATGAACTTTGATCAGGTTTGTGAGTGTGTCAGAG	5638
Qy	3554	GGTGTATCTGGCTTGGCATCTCTTCTTCTAGTGACCTTTAGGACTTCATATCCAG	3613	Qy	4633	CATAGGGAACACAGGCTTTTCTTACCAACTCAAGGAATTTATCAACTCTGCAACACTTG	4692
Db	4560	GGTGTATCTGGCTTGGCATCTCTTCTTCTAGTGACCTTTAGGACTTCATATCCAG	4619	Db	5639	CATAGGGAACACAGGCTTTTCTTACCAACTCAAGGAATTTATCAACTCTGCAACACTTG	5698
Qy	3614	GTTTCTCTCCACTCGTCCAAAGTCACACGATCTTGGCACTCTAACTTAAATGCAAAATA	3673	Qy	4693	CGTATGAGGTAGCAAGGAAATGTCATATCTTGAAGTCGACAGTGTAGTCTTTGAG	4752
Db	4620	GTTTCTCTCCACTCGTCCAAAGTCACACGATCTTGGCACTCTAACTTAAATGCAAAATA	4679	Db	5699	CGTATGAGGTAGCAAGGAAATGTCATATCTTGAAGTCGACAGTGTAGTCTTTGAG	5758
Qy	3674	AAATAGTCAGACATCTCCAGGCTATCTTCTTGGATTTAGCTTCTGCAAGTTTATC	3733	Qy	4753	AAATTTCTGAAGCCGTAATTTTATTTATCAGTGTGATCTATCAGAGATCTCTACGCCG	4812
Db	4680	AAATAGTCAGACATCTCCAGGCTATCTTCTTGGATTTAGCTTCTGCAAGTTTATC	4739	Db	5759	AAATTTCTGAAGCCGTAATTTTATTTATCAGTGTGATCTATCAGAGATCTCTACGCCG	5818
Qy	3734	AGCTTCTCTCTAAATTTAGCGTTC-AACAAAACCTTCTGCTGCTCAAAATAACCGTTTGGTAT	3792	Qy	4813	GACGATCTGGCGGAGCTGCGGCTGCGATCACCGGCGCACAGGTCGCTTGTGCGG	4872
Db	4740	AGCTTCTCTCTAAATTTAGCGTTC-AACAAAACCTTCTGCTGCTCAAAATAACCGTTTGGTAT	4799	Db	5819	GACGATCTGGCGGAGCTGCGGCTGCGATCACCGGCGCACAGGTCGCTTGTGCGG	5878
Qy	3793	AAGAACTTCTGAGGATCTCTTACGATCCCAAGGTTGGTCTCAGTGGCTCTAAGAC	3852	Qy	4873	CCTATATCGCCGACATCAACCGATGGGAAGATCGGGTCTGCACTTCTGGGCTCATGAGCG	4932
				Db	5879	CCTATATCGCCGACATCAACCGATGGGAAGATCGGGTCTGCACTTCTGGGCTCATGAGCG	5938

8099 CGCGTTTCGGTGATGACGGTGAACACCTCTTGACACATGACGCTCCCGAGACGGTCAAG 8158
5913 CTTGCTCTGAAGCGGATCCCGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGTTG 5972
8159 CTTGCTCTGAAGCGGATCCCGGAGCAGACAAGCCCGTCAGGGCGCGTCAAGCGGTTG 8218
5973 CGGGGTGCGGGCGCAGCATGACCCAGTCACCTGAGCGATGAGCGAGTGATGATGCT 6032
8219 CGGGGTGCGGGCGCAGCATGACCCAGTCACCTGAGCGATGAGCGAGTGATGATGCT 8278
6033 TAACTATGCGGCATCAGAGCAGATGTTACTGAGAGTGACCATATGCGGTTGGAATACC 6092
8279 TAACTATGCGGCATCAGAGCAGATGTTACTGAGAGTGACCATATGCGGTTGGAATACC 8338
6093 GCACAGATGCTGAGGAGAAATACCGCATCAGGGGCTCTCCGTTCTCGCTCACTGA 6152
8339 GCACAGATGCTGAGGAGAAATACCGCATCAGGGGCTCTCCGTTCTCGCTCACTGA 8398
6153 CTCGCTCGGCTCGGTCGCTCGGTCGCGGAGCGGTATCAGCTCACTCAAAAGCGGTAAT 6212
8399 CTCGCTCGGCTCGGTCGCTCGGTCGCGGAGCGGTATCAGCTCACTCAAAAGCGGTAAT 8458
6213 ACGGTTATCCACAGAAATCAGGGATACCGGAGAAAGAACATGTGAGCAAAAGCCAGCA 6272
8459 ACGGTTATCCACAGAAATCAGGGATACCGGAGAAAGAACATGTGAGCAAAAGCCAGCA 8518
6273 AAAGCCAGGACCGTAAAGGCGCGTGTGCGGTTTCCATAGGCTCCGCGCCGCC 6332
8519 AAAGCCAGGACCGTAAAGGCGCGTGTGCGGTTTCCATAGGCTCCGCGCCGCC 8578
6333 TCACGAGCATCAAAAAATCAGCGCTCAAGTCAGAGGTGGCGAAACCCGACAGCACTATA 6392
8579 TCACGAGCATCAAAAAATCAGCGCTCAAGTCAGAGGTGGCGAAACCCGACAGCACTATA 8638
6393 AAGATACAGGGGTTTCCCGCTGGAGGCTCCCTCGTGGCTCTCTGTCGACCTGCC 6452
8639 AAGATACAGGGGTTTCCCGCTGGAGGCTCCCTCGTGGCTCTCTGTCGACCTGCC 8698
6453 GCTTACCGGATACCTGTCGCGCTTCTCCCTTCGCGAAGCGTGGCGCTTCTCATAGTCTC 6512
8699 GCTTACCGGATACCTGTCGCGCTTCTCCCTTCGCGAAGCGTGGCGCTTCTCATAGTCTC 8758
6513 ACGCTGTAGGTATCTCAGTTCGGTGTGAGTCTGCTCCAGTGGGCTGTGTCACGA 6572
8759 ACGCTGTAGGTATCTCAGTTCGGTGTGAGTCTGCTCCAGTGGGCTGTGTCACGA 8818
6573 ACCCCCGTTACCGCGGACCGCTGCGCTTATCCGCTAACTATCGTTCGATCCAAACC 6632
8819 ACCCCCGTTACCGCGGACCGCTGCGCTTATCCGCTAACTATCGTTCGATCCAAACC 8878
6633 GGTAAACACGACTTATCGCCACTGSCAGCAGCCACTGGTAAACAGATTTAGCAGCGAG 6692
8879 GGTAAACACGACTTATCGCCACTGSCAGCAGCCACTGGTAAACAGATTTAGCAGCGAG 8938
6693 GTATGTAGGGGTGCTACAGAGTCTTGAAGTGGTGGCTTAAGTGGTGGTGGTGGTGGT 6752
8939 GTATGTAGGGGTGCTACAGAGTCTTGAAGTGGTGGCTTAAGTGGTGGTGGTGGTGGT 8998
6753 GACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTACCTTCGGAAGAGGTTGGTAG 6812
8999 GACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTACCTTCGGAAGAGGTTGGTAG 9058
6813 CTTTGTATCGGCAAAACAAACCCAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 6872
9059 CTTTGTATCGGCAAAACAAACCCAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 9118
6873 GATTACCGCGAGAAAGAGTCTCAGAGATCTTGTGATCTTTCTACGGGTCTGA 6932
9119 GATTACCGCGAGAAAGAGTCTCAGAGATCTTGTGATCTTTGTGATCTTTCTACGGGTCTGA 9178
6933 CGCTCAGTGGAACGAAACCTCAGGTTAAGGGATTTTGGTCATGAGATTTCAAAAGGAT 6992

9179 CGCTCAGTGGAAACAACTCAGCTTAAGGGATTTTGGTCATGAGATTTCAAAAGGAT 9238
5993 CTTCACTAGATCCCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAGTATATAGA 7052
9239 CTTCACTAGATCCCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAGTATATAGA 9298
7053 GTAAACTTGGTCTGACAGATTACCAATGCTTAATCAGTGAAGCACTTATCTCAGCGATCTG 7112
9299 GTAAACTTGGTCTGACAGATTACCAATGCTTAATCAGTGAAGCACTTATCTCAGCGATCTG 9358
7113 TCTATTTGGTCTCATCATAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7172
9359 TCTATTTGGTCTCATCATAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9418
7173 GGGCTTACCATCTGGGCCCGCAGTGTGCAATGATACCGCGAGACCCACGCTCACC GGCTCC 7232
9419 GGGCTTACCATCTGGGCCCGCAGTGTGCAATGATACCGCGAGACCCACGCTCACC GGCTCC 9478
7233 AGATTTATCAGCAATTAACACAGCAGCGAGGCGCGAGCGAGGAGTGGTCTGCAAC 7292
9479 AGATTTATCAGCAATTAACACAGCAGCGAGGCGCGAGCGAGGAGTGGTCTGCAAC 9538
7293 TTTATCGGCTCCATCCAGTCTATTAAATTTGTCGGGAGAGCTAGAGTAAGTAGTTCGCC 7352
9539 TTTATCGGCTCCATCCAGTCTATTAAATTTGTCGGGAGAGCTAGAGTAAGTAGTTCGCC 9598
7353 AGTTAATAGTTTGGCAACGCTTCCCATTTGTCGGGAGAGCTAGAGTAAGTAGTTCGCC 7412
9599 AGTTAATAGTTTGGCAACGCTTCCCATTTGTCGGGAGAGCTAGAGTAAGTAGTTCGCC 9658
7413 GTTTGTATGGCTTCACTCAGCTCCCGTCCCAACCATCAAGCGAGTGTACATGATCCCC 7472
9659 GTTTGTATGGCTTCACTCAGCTCCCGTCCCAACCATCAAGCGAGTGTACATGATCCCC 9718
7473 CATGTTGTGCAAAAAAGCGGTAGCTCTTCCGTCCTCCGATCGTTGTTCAGAAAGTAGTT 7532
9719 CATGTTGTGCAAAAAAGCGGTAGCTCTTCCGTCCTCCGATCGTTGTTCAGAAAGTAGTT 9778
7533 GGGCGAGTGTATCACTCATGTGTTAGCAGCAGTGCATTAATCTTCTACTCTCATGCC 7592
9779 GGGCGAGTGTATCACTCATGTGTTAGCAGCAGTGCATTAATCTTCTACTGTCAATGCC 9838
7593 ATCCGTAAAGTGTCTTCTGTCGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 7652
9839 ATCCGTAAAGTGTCTTCTGTCGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 9898
7653 TATCGCGGACCGAGTTGCTTCTTCCCGGCGTCAACACGGGATAATACCGCGCACATAG 7712
9899 TATCGCGGACCGAGTTGCTTCTTCCCGGCGTCAACACGGGATAATACCGCGCACATAG 9958
7713 CAGAACTTTAAAGTGTCTCATTTGGAACACGTTCTTCCGGGCGAAACTCTTCAAGCAT 7772
9959 CAGAACTTTAAAGTGTCTCATTTGGAACACGTTCTTCCGGGCGAAACTCTTCAAGCAT 10018
7773 CTTACCGCTGTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCCACTGATCTTCAAG 7832
10019 CTTACCGCTGTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCCACTGATCTTCAAG 10078
7833 ATCTTTTACTTTACCGAGCTTCTGAGTGAACACGTTCTTCCGGGCGAAACTCTTCAAGCAT 7892
10079 ATCTTTTACTTTACCGAGCTTCTGAGTGAACACGTTCTTCCGGGCGAAACTCTTCAAGCAT 10138
7893 AAAGGGAATAGGCGGACACGGAATGTTGAATCTCATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 7952
10139 AAAGGGAATAGGCGGACACGGAATGTTGAATCTCATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 10198
7953 TTGAAGCATTTATCAGGGTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAA 8012
10199 TTGAAGCATTTATCAGGGTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAA 10258
8013 AAATAACCAATAGGGTTCGCGCAGATTTCCCGGAAAGTGCACCTGACGCTAAGA 8072
10259 AAATAACCAATAGGGTTCGCGCAGATTTCCCGGAAAGTGCACCTGACGCTAAGA 10318

```
QY      8073 AACCATTTATCATGACATTAACCTATAAAATAGGCGTATCAGAGGCCCTTTTGGTCT 8132
Db      10319 AACCATTTATCATGACATTAACCTATAAAATAGGCGTATCAGAGGCCCTTTTGGTCT 10378
QY      8133 TCAAGAAATAATTCATGTTTGCAGCTTATCATCGATAAGCTGACATCATGTTGGTATT 8192
Db      10379 TCAAGAAATAATTCATGTTTGCAGCTTATCATCGATAAGCTGACATCATGTTGGTATT 10438
QY      8193 GTGAATAGACGACATCGGGAACACTGAAATAAACAAGTTATTATTG 8241
Db      10439 GTGAATAGACGACATCGGGAACACTGAAATAAACAAGTTATTATTG 10487

RESULT 15
AX417107
LOCUS      AX417107 10491 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 32 from Patent WO0198344.
ACCESSION AX417107
VERSION    AX417107.1 GI:21445694
KEYWORDS   unidentified
SOURCE      unidentified
ORGANISM    unidentified
REFERENCE   1
AUTHORS     Ling,L.E. and Sanicola-Nadel,M.
TITLE       Angiogenesis-modulating compositions and uses
JOURNAL     Patent: WO 0198344-A 32 27-DEC-2001;
            BIOGEN, INC. (US)
FEATURES    Location/Qualifiers
             1..10491
             /organism="unidentified"
             /mol_type="unassigned DNA"
             /db_xref="taxon:32644"
             /note="Plasmid pUB114"

ORIGIN
Query Match 68.8%; Score 5665.8; DB 6; Length 10491;
Best Local Similarity 77.2%; Pred. No. 0;
Matches 8100; Conservative 0; Mismatches 137; Indels 2252; Gaps 10;

QY      2 GATCTAAATCCAAAGACGAAAGTTGAATGAACCTTTTGCCATCCGACATCCACAGG 61
Db      1 GATCTAAATCCAAAGACGAAAGTTGAATGAACCTTTTGCCATCCGACATCCACAGG 60
QY      62 TCCATTCTCACACATAAGTGCACAAACGCAACGAGGAGGATACACTAGCAGCAGACCGTT 121
Db      61 TCCATTCTCACACATAAGTGCACAAACGCAACGAGGAGGATACACTAGCAGCAGACCGTT 120
QY      122 GCAACGAGACGACCTCCATCTCTCTCTCTCTCAACACCCACTTTGGCATCGAATAACCA 181
Db      121 GCAACGAGACGACCTCCATCTCTCTCTCTCTCAACACCCACTTTGGCATCGAATAACCA 180
QY      182 GCCAGTTATTGGGCTTCATTTGGAGCTCGCTCATTTCCAAATCTCTATTAGGCTACTAA 241
Db      181 GCCAGTTATTGGGCTTCATTTGGAGCTCGCTCATTTCCAAATCTCTATTAGGCTACTAA 240
QY      242 CACCATGACTTTATPAGCTGTCTATCTGCTGCCCCCTTGGCGAGGTTCAATGTTGTTTAT 301
Db      241 CACCATGACTTTATPAGCTGTCTATCTGCTGCCCCCTTGGCGAGGTTCAATGTTGTTTAT 300
QY      302 TTCCGAATGCAACAGCTCCGCAATTAACCCGCAACATCACTCCAGATGAGGCTTTCTGA 361
Db      301 TTCCGAATGCAACAGCTCCGCAATTAACCCGCAACATCACTCCAGATGAGGCTTTCTGA 360
QY      362 GTGTGGGTCAAAATAGTTTCATGTTCCCAATGGCCAAATCGACAGTTTAAACGCTG 421
Db      361 GTGTGGGTCAAAATAGTTTCATGTTCCCAATGGCCAAATCGACAGTTTAAACGCTG 420
QY      422 TCTTGGAACTTAATGACAAAGCGGTGATCTCATCCAAAGTGAATGATGTTGGTTCGT 481
Db      421 TCTTGGAACTTAATGACAAAGCGGTGATCTCATCCAAAGTGAATGATGTTGGTTCGT 480
```

```
QY      482 TGAATGCTAAGCGCAGTTGGTCAAAAGAAACCTTCCAAAGTCCCATACCGTTTGTTC 541
Db      481 TGAATGCTAAGCGCAGTTGGTCAAAAGAAACCTTCCAAAGTCCCATACCGTTTGTTC 540
QY      542 TTGTTTGGTATTGATTCAGAAATGCTCAAAAATAATCTCATTAATGCTTAGCGCAGTCTC 601
Db      541 TTGTTTGGTATTGATTCAGAAATGCTCAAAAATAATCTCATTAATGCTTAGCGCAGTCTC 600
QY      602 TCTATCGCTTCTGAACCCCGGTGCACTGCTGCGGAAAACGCAAAATGGGAAACACCGCTT 661
Db      601 TCTATCGCTTCTGAACCCCGGTGCACTGCTGCGGAAAACGCAAAATGGGAAACACCGCTT 660
QY      662 TTGATGATGATTATGCAATGCTCCACATGTAATGCTTCAAGATCTCTGGTGGGAATCTG 721
Db      661 TTGATGATGATTATGCAATGCTCCACATGTAATGCTTCAAGATCTCTGGTGGGAATCTG 720
QY      722 CTGATAGCCTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTGCAGCAATA 781
Db      721 CTGATAGCCTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTGCAGCAATA 780
QY      782 TATAACAGAGGAAAGCTGCCCTGCTTAAACCTTTTTTTTATCATCATATTAGCTTA 841
Db      781 TATAACAGAGGAAAGCTGCCCTGCTTAAACCTTTTTTTTATCATCATATTAGCTTA 839
QY      842 CTTTCATAATTGCGACTGGTTCCAATTCAGACAGCTTTTGAATTTTAAACGCTTTTAAACGAC 901
Db      840 CTTTCATAATTGCGACTGGTTCCAATTCAGACAGCTTTTGAATTTTAAACGCTTTTAAACGAC 899
QY      902 AACTTGAGAGATCAAAAACAACTAAATTTTCGAGGATCCAAACGATGAGATTCCTT 961
Db      900 AACTTGAGAGATCAAAAACAACTAAATTTTCGAGGATCCAAACGATGAGATTCCTT 959
QY      962 CAATTTTACTCAGCTTTTATTTGCGAGCATCTCCGCAATTAGCTCTCCAGTCAACACTA 1021
Db      960 CAATTTTACTCAGCTTTTATTTGCGAGCATCTCCGCAATTAGCTCTCCAGTCAACACTA 1019
QY      1022 CAACAGAGATGAAACGCGCAAAATTCGGCTGAAGCTGTCAATCGGTTACTCAGATTAG 1081
Db      1020 CAACAGAGATGAAACGCGCAAAATTCGGCTGAAGCTGTCAATCGGTTACTCAGATTAG 1079
QY      1082 AAGGGGATTTCCATGTTGCTGTTTTCATTTTCCAAACAGCACAATTAACGGTTATTGT 1141
Db      1080 AAGGGGATTTCCATGTTGCTGTTTTCATTTTCCAAACAGCACAATTAACGGTTATTGT 1139
QY      1142 TTATAAATACTACTATPCCAGCATGCTGCTPAAAGAAAGAGGGGTATCTCTCGAGAAAA 1201
Db      1140 TTATAAATACTACTATPCCAGCATGCTGCTPAAAGAAAGAGGGGTATCTCTCGAGAAAA 1199
QY      1202 GAGAGCTGAAGCCCGAGG----- 1219
Db      1200 GATGCGACCGGCGAGGGGTTGCGGAAAGAGAGGACGCCCAAAAGCTGACCCCTTTAG 1259
QY      1220 -----CCAGGCCAGGCCCGCCAGCCAG 1242
Db      1260 CCTACAAGCAGTTTATCCCAATGTGGCGGAGAGACCTTAGCGCCAGCGAGGATG 1319
QY      1243 GAAGAAACAT----- 1252
Db      1320 AAGGGAAGATCTCCAGAAACTCCGAGCGATTTAAGAACTCACCCCCAAATTAACCCCG 1379
QY      1253 -----GTATCATGGCCCCCGGAGAGGATAAATTTGGTGTCTCCCGGTGTC 1299
Db      1380 ACATCATATTTAAGGATGAAGAAACACCGGAGCGGACAGGCTGATGACTCAGAGGTCTA 1439
QY      1300 A----- 1300
Db      1440 AGGACAAGTTGAACGCTTTGGCCATCTCGGTATGAACAGTCAGTGCGCAGGATGAACCTGC 1499
QY      1301 ----- 1300
Db      1500 GGTGACCGAGGGCTGGGACGAGATGSCCACCACCTCAGAGGATCTCTGCATACGAGG 1559
QY      1301 ----- 1301
```


Db 1560 GCCGCGCAGTGGACATCACCAGCTCTGACCGGACCGCAGCAAGTACCGGCATGCTGGCCC 1619
Qy 1301 ----- 1300
Db 1620 GCTGGCGGTGGAGCGCGCTTCGACTGGGTGTACTACGAGTCCAAAGCACAATATCCACT 1679
Qy 1301 ----- 1300
Db 1680 GCTCGGTGAACGACAGAACTCGGTGGCGGCCAAATCGGAGGCGTGCACGTGCCAGGG 1739
Qy 1301 ----- 1300
Db 1740 ATTGTGGTTGTAAGCCTTCATATGTACAGTCCAGAGATATCTCTCTTCATCTCC 1799
Qy 1301 ----- 1300
Db 1800 CCCCAAAGCCAAAGGATGTGCTACCAATTACTCTGACTTAAGTCACTGCTGTGTGG 1859
Qy 1301 ----- 1300
Db 1860 TAGACATCAGCAAGGATGATCCGAGGTCCTCAGCTGGTTTGTAGATGATGTGGAGG 1919
Qy 1301 ----- 1300
Db 1920 TGCACACAGCTCAGACGCAACCAACGGAAGACAGTTCCAAAGCACTTTCCGCTCAGTCA 1979
Qy 1301 ----- 1300
Db 1980 GTGAATCCCATCATGCACAGGACTGGCTCAATGGCAAGAGTTCAAATGCAGGGTCA 2039
Qy 1301 ----- 1300
Db 2040 ACAGTGCAGCTTTCCTGCGCCCATCGAGAAACCATCTCCAAACCAAGGACAGACCGA 2099
Qy 1301 ----- -CG 1303
Db 2100 AGGCTCCACAGGTGTACACCTTCACCTCCCAAGAGCAGATGCCAAGATTAAGTCA 2159
Qy 1304 CCCAGCAGTGCACGAGAGAGGTGTCTGTTTGTATGACAGTGTCCGGGGATTTCCCGTGGT 1363
Db 2160 GTCTGACCTGCTGATTAACAGACTTCTTCCCTGGAAGACATTACTGTGGAGTGGCAGTGA 2219
Qy 1364 GCTTCCACCCCATGGCCATCGAGACA ----- 1390
Db 2220 ATGGGCAGCAGCGGAGAACTCAAGAACACTCAGCCCATCATGGACACAGATGGCTCTT 2279
Qy 1391 ----- 1390
Db 2280 ACTTCGCTACAGCAAGCTCAATGTGCAGAGCAACTGGGAGGACGGAATACATTCA 2339
Qy 1391 ----- CTAAGAGAGAAATGTCCCTTCTAAT 1418
Db 2340 CCTGCTCTGTATACATGAGGCGCTGCACAAACCATACTGAGAAGAGCCTCTCCACT 2399
Qy 1419 AGTGGGTAGA ----- ATTCCCTAGGCGCGCGCAATTA 1453
Db 2400 CTCCTGGTAAATGATCCCAAGTCTCTGGAGCCCTCTGTGCTCTCAGCGGCGCGAAATTA 2459
Qy 1454 ATTCGCTTTAGACATGACTGTTCTCAGTTCAGTTGGGCACCTTACGAGAAGACCGGTCT 1513
Db 2460 ATTCGCTTTAGACATGACTGTTCTCAGTTCAGTTGGGCACCTTACGAGAAGACCGGTCT 2519
Qy 1514 TGTAGATTTCAATCAAGAGGATGTGCAATGCCATTTGCCCTGAGAGATGAGGCTTCAT 1573
Db 2520 TGTAGATTTCAATCAAGAGGATGTGCAATGCCATTTGCCCTGAGAGATGAGGCTTCAT 2579
Qy 1574 TTTTGTATCTTTTATTTGTAACCTATATAGTATAGGATTTTTTTTGTCAATTTTGTTC 1633
Db 2580 TTTTGTATCTTTTATTTGTAACCTATATAGTATAGGATTTTTTTTGTCAATTTTGTTC 2639
Qy 1634 TTTCTGTACAGCTTCTCTGTATCAGCCTATCTCCAGCTGATGAATCTTTGTGGTAG 1693

Db 2640 TTCTCGTACGAGCTTCTCTGTATCAGCCTATCTCGCAGCTGATGAATATCTTTGTGGTAG 2699
Qy 1694 GGGTTTGGGAAATCAATTCGAGTTTCTGATGTTTTCTTGTATTTCCACTCCCTCTCAGA 1753
Db 2700 GGGTTTGGGAAATCAATTCGAGTTTCTGATGTTTTCTTGGTATTTCCACTCTCTTCAGA 2759
Qy 1754 GTACAGAAATTAAGTGAAGTTCGTTTGTGCAAGCTTATCGATAAGCTTTAATCGGT 1813
Db 2760 GTACAGAAATTAAGTGAAGTTCGTTTGTGCAAGCTTATCGATAAGCTTTAATCGGT 2819
Qy 1814 AGTTTATCAGATTAATTTCTTAACGAGTCAGGCAACCGTATGAATTAACATCAATGCG 1873
Db 2820 AGTTTATCAGATTAATTTCTTAACGAGTCAGGCAACCGTATGAATTAACATCAATGCG 2879
Qy 1874 CTGATCGTATCTCCGCAACCGTCACCTCGATGCTGTAGGATAGGCTTTGTTATGCCG 1933
Db 2880 CTGATCGTATCTCCGCAACCGTCACCTCGATGCTGTAGGATAGGCTTTGTTATGCCG 2939
Qy 1934 GTACTGCGGCGCTCTTTGCGGATATCGTCCATTCGCAAGCATCGCAGTCACTATGCG 1993
Db 2940 GTACTGCGGCGCTCTTTGCGGATATCGTCCATTCGCAAGCATCGCAGTCACTATGCG 2999
Qy 1994 GTGCTGTAGCGCTATATGATGCAATTTCTATGCGCACCGTTCCTCGAGCACTG 2053
Db 3000 GTGCTGTAGCGCTATATGATGCAATTTCTATGCGCACCGTTCCTCGAGCACTG 3059
Qy 2054 TCCGACCGCTTTGGCGCGCCCGCAGTCTGCTCGTCTGCTACTTGGAGCACTATCGAC 2113
Db 3060 TCCGACCGCTTTGGCGCGCCCGCAGTCTGCTCGTCTGCTACTTGGAGCACTATCGAC 3119
Qy 2114 TAGCGATCATGCGGACCAACCGGCTCTGCTGATCTATCGAATCTAAATGTAAGTTAA 2173
Db 3120 TAGCGATCATGCGGACCAACCGGCTCTGCTGATCTATCGAATCTAAATGTAAGTTAA 3179
Qy 2174 ATCTCTAAATAATTAATAAGTCCAGTTCCTCATACGAACTTAAACAGCATTTGGCGTG 2233
Db 3180 ATCTCTAAATAATTAATAAGTCCAGTTCCTCATACGAACTTAAACAGCATTTGGCGTG 3239
Qy 2234 AGCATCTAGACCTTCAACAGCAGCAGATCCATCTGCTTGGCCATATGTTTCAGTCC 2293
Db 3240 AGCATCTAGACCTTCAACAGCAGCAGATCCATCTGCTTGGCCATATGTTTCAGTCC 3299
Qy 2294 CTGAGGAGTTTACGTTCTGTAAGTGAATCTGTAAGTGTGCACTTAACTCCGCT 2353
Db 3300 CTGAGGAGTTTACGTTCTGTAAGTGAATCTGTAAGTGTGCACTTAACTCCGCT 3359
Qy 2354 GTATTGACGGGCATATCCGTACGTTGGCAAAGTGTGGTACCGGAGAGTAATCTCC 2413
Db 3360 GTATTGACGGGCATATCCGTACGTTGGCAAAGTGTGGTACCGGAGAGTAATCTCC 3419
Qy 2414 ACAACTCTCTGGAGAGTAGGACCAACAAACACAGATCCAGCGTGTGTTGATCAAC 2473
Db 3420 ACAACTCTCTGGAGAGTAGGACCAACAAACACAGATCCAGCGTGTGTTGATCAAC 3479
Qy 2474 ATAGAGAGAGCAATTCGATTTGCAAGTGTTCAGGAGCGTACTGATTTGACAT 2533
Db 3480 ATAGAGAGAGCAATTCGATTTGCAAGTGTTCAGGAGCGTACTGATTTGACAT 3539
Qy 2534 TTCCAAAGCCCTGCTGATGTTGCAACCGATAGGTTGTAGAGTGTCAATACACTTGGC 2593
Db 3540 TTCCAAAGCCCTGCTGATGTTGCAACCGATAGGTTGTAGAGTGTCAATACACTTGGC 3599
Qy 2594 TACAATTTCAACCTTGGCAACTGCAAGCTTGGTGTGAACAGCACTTCAATTTGGC 2653
Db 3600 TACAATTTCAACCTTGGCAACTGCAAGCTTGGTGTGAACAGCACTTCAATTTGGC 3659
Qy 2654 AAGTCCTTGTCTGATATCGACAGCAACAGATCACCTGGGAATCAATACCATGTC 2713
Db 3660 AAGTCCTTGTCTGATATCGACAGCAACAGATCACCTGGGAATCAATACCATGTC 3719
Qy 2714 AGCTTGACAGAGGCTCGAGCAACAGATCTGGAATCAGCGTATTTATCAGCAATAAC 2773
Db 3720 AGCTTGAG-CAGAGGCTCGAGCAACAGATCTGGAATCAGCGTATTTATCAGCAATAAC 3778

QY	2774	TAGAACTT	CAGAAAGCC	CAGCAGCAT	GTCAA	TACTAC	CAGGCG	CTGAT	GTGTCAT	TTTG	2833				
Db	3779	TAGAACTT	CAGAAAGCC	CAGCAGCAT	GTCAA	TACTAC	CAGGCG	TGAT	GTGTCAT	TTTG	3838				
QY	2834	AACGATCAT	CTTGGCAGCAG	TAAAGCA	ACTGG	TTCTTG	AGACCA	AAATAT	TTTGTGCAC	ACTT	2893				
Db	3839	AACGATCAT	CTTGGCAGCAG	TAAAGCA	ACTGG	TTCTTG	AGACCA	AAATAT	TTTGTGCAC	ACTT	3898				
QY	2894	AGAAACAG	TTTCTG	TTCCG	TGAAGC	CATAG	CAGTACT	CGCTTGG	CGCGCT	CCTGCTAG	CAC	2953			
Db	3899	AGAAACAG	TTTCTG	TTCCG	TGAAGC	CATAG	CAGTACT	CGCTTGG	CGCGCT	CCTGCTAG	CAC	3958			
QY	2954	GATACACT	TAGCACC	AACTTT	TGGGCA	ACGTAG	ATGAC	TTCTTGGG	TAGGGT	TAACCAT	C	3013			
Db	3959	GATACACT	TAGCACC	AACTTT	TGGGCA	ACGTAG	ATGAC	TTCTTGGG	TAGGGT	TAACCAT	C	4018			
QY	3014	CTTCTTAG	TGGAGAT	CAAAAA	CAAA	TTTCTT	TGGAC	CAGCA	ACTTTTGG	CAGGAA	CACC	3073			
Db	4019	CTTCTTAG	TGGAGAT	CAAAAA	CAAA	TTTCTT	TGGAC	CAGCA	ACTTTTGG	CAGGAA	CACC	4078			
QY	3074	CAGGATC	AGGGAAG	TGGA	AGGC	GAATTTG	CGGTTC	CAC	CAGGAT	ATATAG	AGGCCA	ACTTT	3133		
Db	4079	CAGGATC	AGGGAAG	TGGA	AGGC	GAATTTG	CGGTTC	CAC	CAGGAT	ATATAG	AGGCCA	ACTTT	4138		
QY	3134	CTCAATAG	GTCTTTG	CAAAA	CGAG	AGCAG	ACTAC	CAC	AGGCGCA	AGTCTCA	CACTTGC	AAACGT	3193		
Db	4139	CTCAATAG	GTCTTTG	CAAAA	CGAG	AGCAG	ACTAC	CAC	AGGCGCA	AGTCTCA	CACTTGC	AAACGT	4198		
QY	3194	CTCGTTAG	TGAGCTT	CA	TGGAA	TTTCTG	ATCG	TTATCT	ATAG	AGAGAT	CAATGG	CTCT	3253		
Db	4199	CTCGTTAG	TGAGCTT	CA	TGGAA	TTTCTG	ATCG	TTATCT	ATAG	AGAGAT	CAATGG	CTCT	4258		
QY	3254	CTTAAAG	CGTATCTG	GC	CAATTC	GCA	TAACTCTCTG	GGAAG	AGAGCTTCT	TAA	CAC	AGGTGT	3313		
Db	4259	CTTAAAG	CGTATCTG	GC	CAATTC	GCA	TAACTCTCTG	GGAAG	AGAGCTTCT	TAA	CAC	AGGTGT	4318		
QY	3314	CTTCAAA	AGGCACTT	CCAT	CAAA	CTTGG	CAGTTAG	TTCTTAA	AGGCGCTT	TG	CA	CCATTTTG	3373		
Db	4319	CTTCAAA	AGGCACTT	CCAT	CAAA	CTTGG	CAGTTAG	TTCTTAA	AGGCGCTT	TG	CA	CCATTTTG	4378		
QY	3374	ACGAACAT	TGTG	CA	CAATTTG	GTG	ACTAA	TCCAT	AAATCTG	TTCCG	TTTCTG	GATAGG	3433		
Db	4379	ACGAACAT	TGTG	CA	CAATTTG	GTG	ACTAA	TCCAT	AAATCTG	TTCCG	TTTCTG	GATAGG	4438		
QY	3434	ACGACGA	AGGCGAT	CTT	CAATTTCT	TGT	AGAGG	CGCTT	AGAA	ACGTCA	ATTTTGC	ACAA	3493		
Db	4439	ACGACGA	AGGCGAT	CTT	CAATTTCT	TGT	AGAGG	CGCTT	AGAA	ACGTCA	ATTTTGC	ACAA	4498		
QY	3494	TTCAAT	ACGACCTT	CAGA	AGGCA	CTTCTT	TAG	GTTTG	GATTTCTT	T	TAG	TTGTTCTCCT	3553		
Db	4499	TTCAAT	ACGACCTT	CAGA	AGGCA	CTTCTT	TAG	GTTTG	GATTTCTT	T	TAG	TTGTTCTCCT	4558		
QY	3554	GGTGTA	TCCTGG	CTTGG	CACTC	CTCTT	CTTCTAG	TG	ACCTTT	TAG	GCACTTCA	TATCCAG	3613		
Db	4559	GGTGTA	TCCTGG	CTTGG	CACTC	CTCTT	CTTCTAG	TG	ACCTTT	TAG	GCACTTCA	TATCCAG	4618		
QY	3614	GTTTCT	CTCCAC	CTCG	TCCAA	AGTC	CAC	ACG	TACTTGG	CACTCTAA	CTAATG	CAAAATA	3673		
Db	4619	GTTTCT	CTCCAC	CTCG	TCCAA	AGTC	CAC	ACG	TACTTGG	CACTCTAA	CTAATG	CAAAATA	4678		
QY	3674	AAATAAG	TAGC	ACATTTCC	CA	AGGCTAT	CTTCTT	CGGA	TTT	TAG	CTTCTG	CAAGTTCA	CTC	3733	
Db	4679	AAATAAG	TAGC	ACATTTCC	CA	AGGCTAT	CTTCTT	CGGA	TTT	TAG	CTTCTG	CAAGTTCA	CTC	4738	
QY	3734	AGCTT	CTCCT	CTAATTTT	TAG	CGTTT	C	CAAA	AACTCG	T	CGT	CAAAATA	CGGTTTGG	3792	
Db	4739	AGCTT	CTCCT	CTAATTTT	TAG	CGTTT	C	CAAA	AACTCG	T	CGT	CAAAATA	CGGTTTGG	4798	
QY	3793	AGAACTT	CTG	AG	CAATTTG	CTT	ATCG	ATCC	CA	AGGTTG	GGCTT	CAATGG	CTCTA	AGAC	3852
Db	4799	AGAACTT	CTG	AG	CAATTTG	CTT	ATCG	ATCC	CA	AGGTTG	GGCTT	CAATGG	CTCTA	AGAC	4857

3853	Qy	CTTTTGATTGGCCAAACACAGGAAGTCGGTTCCAAAGTGACAGAAAACCAACACCTGTTTGT	3912	Qy	TTTGTGCTTATAGAGCTCGGNAATAGACTTTTGGACGAGTACACAGGCCCAACGAGCT
4858	Db	CCTTTGATTGGCCAAACACAGGAAGTCGGTTCCAAAGTGACAGAAAACCAACACCTGTTTGT	4917	Qy	CAACCAAAATTTCAAGCAGTCTCAATCAAAATCCAAATTCGATACCCAGCAACTTTTGAG
3913	Qy	CAACCAAAATTTCAAGCAGTCTCAATCAAAATCCAAATTCGATACCCAGCAACTTTTGAG	4977	Db	CAACCAAAATTTCAAGCAGTCTCAATCAAAATTCGATACCCAGCAACTTTTGAG
4918	Db	CAACCAAAATTTCAAGCAGTCTCAATCAAAATTCGATACCCAGCAACTTTTGAG	3973	Qy	TTGCTCCAGATGAGCACCTTTATACCAACAAACCGTGACGACGAGATGGTAGATCCAG
3973	Qy	TTGCTCCAGATGAGCACCTTTATACCAACAAACCGTGACGACGAGATGGTAGATCCAG	4032	Qy	TTTGTGCTTATAGAGCTCGGNAATAGACTTTTGGACGAGTACACAGGCCCAACGAGCT
4978	Db	TTCGTCAGATGAGCACCTTTATACCAACAAACCGTGACGACGAGATGGTAGATCCAG	5037	Db	TTTGTGCTTATAGAGCTCGGNAATAGACTTTTGGACGAGTACACAGGCCCAACGAGCT
4033	Qy	TTTGTGCTTATAGAGCTCGGNAATAGACTTTTGGACGAGTACACAGGCCCAACGAGCT	4092	Qy	TTTGTGCTTATAGAGCTCGGNAATAGACTTTTGGACGAGTACACAGGCCCAACGAGCT
5038	Db	TTTGTGCTTATAGAGCTCGGNAATAGACTTTTGGACGAGTACACAGGCCCAACGAGCT	5097	Db	TTTGTGCTTATAGAGCTCGGNAATAGACTTTTGGACGAGTACACAGGCCCAACGAGCT
4093	Qy	AAATTAGAAGAGTCAGGCCAACAAAGTAGTGAATAGACCAATCGGGGGCGTCAGTAGTCAAA	4152	Qy	AAATTAGAAGAGTCAGGCCAACAAAGTAGTGAATAGACCAATCGGGGGCGTCAGTAGTCAAA
5098	Db	AAATTAGAAGAGTCAGGCCAACAAAGTAGTGAATAGACCAATCGGGGGCGTCAGTAGTCAAA	5157	Db	AAATTAGAAGAGTCAGGCCAACAAAGTAGTGAATAGACCAATCGGGGGCGTCAGTAGTCAAA
4153	Qy	ACGCCAAACAAATTTCACTGACAGGGAACTTTTGTGACATCTTCAGAAAGTTCGTATTCA	4212	Qy	ACGCCAAACAAATTTCACTGACAGGGAACTTTTGTGACATCTTCAGAAAGTTCGTATTCA
5158	Db	ACGCCAAACAAATTTCACTGACAGGGAACTTTTGTGACATCTTCAGAAAGTTCGTATTCA	5217	Db	ACGCCAAACAAATTTCACTGACAGGGAACTTTTGTGACATCTTCAGAAAGTTCGTATTCA
4213	Qy	TAGTCAATTCGCCAGCATCAATTAATGGGANTATACACAGAACCAACAGTGGAGTCAAT	4272	Qy	TAGTCAATTCGCCAGCATCAATTAATGGGANTATACACAGAACCAACAGTGGAGTCAAT
5218	Db	TAGTCAATTCGCCAGCATCAATTAATGGGANTATACACAGAACCAACAGTGGAGTCAAT	5277	Db	TAGTCAATTCGCCAGCATCAATTAATGGGANTATACACAGAACCAACAGTGGAGTCAAT
4273	Qy	CTACCAACTTTGCGGTCTCAGAAAAAGCAATACAGTTCTACTACGCGCATTTAGTGAAAC	4332	Qy	CTACCAACTTTGCGGTCTCAGAAAAAGCAATACAGTTCTACTACGCGCATTTAGTGAAAC
5278	Db	CTACCAACTTTGCGGTCTCAGAAAAAGCAATACAGTTCTACTACGCGCATTTAGTGAAAC	5337	Db	CTACCAACTTTGCGGTCTCAGAAAAAGCAATACAGTTCTACTACGCGCATTTAGTGAAAC
4333	Qy	TTTTTCAAAATCGCCAGTGAGAGAAAGGACACGCGATCTAGCATTTAGCGGGCAAGG	4392	Qy	TTTTTCAAAATCGCCAGTGAGAGAAAGGACACGCGATCTAGCATTTAGCGGGCAAGG
5338	Db	TTTTTCAAAATCGCCAGTGAGAGAAAGGACACGCGATCTAGCATTTAGCGGGCAAGG	5397	Db	TTTTTCAAAATCGCCAGTGAGAGAAAGGACACGCGATCTAGCATTTAGCGGGCAAGG
4393	Qy	ATGCACACTTTATCAACACAGGGTCCTATAGATAACCCCTAGCGCTGGGATCATCCTTTGA	4452	Qy	ATGCACACTTTATCAACACAGGGTCCTATAGATAACCCCTAGCGCTGGGATCATCCTTTGA
5398	Db	ATGCACACTTTATCAACACAGGGTCCTATAGATAACCCCTAGCGCTGGGATCATCCTTTGA	5457	Db	ATGCACACTTTATCAACACAGGGTCCTATAGATAACCCCTAGCGCTGGGATCATCCTTTGA
4453	Qy	CAACTCTTTCTGCGAAATCTAGTCCAAATCACTTCATTTGATACCAATTTATTTGTAACA	4512	Qy	CAACTCTTTCTGCGAAATCTAGTCCAAATCACTTCATTTGATACCAATTTATTTGTAACA
5458	Db	CAACTCTTTCTGCGAAATCTAGTCCAAATCACTTCATTTGATACCAATTTATTTGTAACA	5517	Db	CAACTCTTTCTGCGAAATCTAGTCCAAATCACTTCATTTGATACCAATTTATTTGTAACA
4513	Qy	TGAGCAAGTTGTTCGATCAGTCCCTCAAAATGGTTCCTCTGTAAACGATGACTCAACTTGA	4572	Qy	TGAGCAAGTTGTTCGATCAGTCCCTCAAAATGGTTCCTCTGTAAACGATGACTCAACTTGA
5518	Db	TGAGCAAGTTGTTCGATCAGTCCCTCAAAATGGTTCCTCTGTAAACGATGACTCAACTTGA	5577	Db	TGAGCAAGTTGTTCGATCAGTCCCTCAAAATGGTTCCTCTGTAAACGATGACTCAACTTGA
4573	Qy	CATTAACTTGAAAGTCAGTCTGATGATGAACTTTGATCAGTTGTGACGCTGGTCAAGCAG	4632	Qy	CATTAACTTGAAAGTCAGTCTGATGATGAACTTTGATCAGTTGTGACGCTGGTCAAGCAG
5578	Db	CATTAACTTGAAAGTCAGTCTGATGATGAACTTTGATCAGTTGTGACGCTGGTCAAGCAG	5637	Db	CATTAACTTGAAAGTCAGTCTGATGATGAACTTTGATCAGTTGTGACGCTGGTCAAGCAG
4633	Qy	CATAGGGAACACAGGCTTTTCTTACCAAACTCAAGGAAATTAACAACCTCTGCAACACTTG	4692	Qy	CATAGGGAACACAGGCTTTTCTTACCAAACTCAAGGAAATTAACAACCTCTGCAACACTTG
5638	Db	CATAGGGAACACAGGCTTTTCTTACCAAACTCAAGGAAATTAACAACCTCTGCAACACTTG	5697	Db	CATAGGGAACACAGGCTTTTCTTACCAAACTCAAGGAAATTAACAACCTCTGCAACACTTG
4693	Qy	CGTATGCAAGTAGCAAGGGAATGTCTACTTGAAGTCGACAGTGAAGTGTAGTCTTTGAG	4752	Qy	CGTATGCAAGTAGCAAGGGAATGTCTACTTGAAGTCGACAGTGAAGTGTAGTCTTTGAG
5698	Db	CGTATGCAAGTAGCAAGGGAATGTCTACTTGAAGTCGACAGTGAAGTGTAGTCTTTGAG	5757	Db	CGTATGCAAGTAGCAAGGGAATGTCTACTTGAAGTCGACAGTGAAGTGTAGTCTTTGAG
4753	Qy	AAATTCGAGCCGATTTTTTATATCAGTGATGATCAGTCAATCAGGAGATCCTCTACGCCG	4812	Qy	AAATTCGAGCCGATTTTTTATATCAGTGATGATCAGTCAATCAGGAGATCCTCTACGCCG
5758	Db	AAATTCGAGCCGATTTTTTATATCAGTGATGATCAGTCAATCAGGAGATCCTCTACGCCG	5817	Db	AAATTCGAGCCGATTTTTTATATCAGTGATGATCAGTCAATCAGGAGATCCTCTACGCCG
4813	Qy	GAGCGATCTGTGGCGACCTGACGCTCGGCATCACCGCGGCCACAGTGGGTTGCTGGCG	4872	Qy	GAGCGATCTGTGGCGACCTGACGCTCGGCATCACCGCGGCCACAGTGGGTTGCTGGCG
5818	Db	GAGCGATCTGTGGCGACCTGACGCTCGGCATCACCGCGGCCACAGTGGGTTGCTGGCG	5877	Db	GAGCGATCTGTGGCGACCTGACGCTCGGCATCACCGCGGCCACAGTGGGTTGCTGGCG
4873	Qy	CCTATATCGCCGACATCACCGATGGGGAAGATCGGGCTCGGCATCTTCGGGCTCATGAGCG	4932	Qy	CCTATATCGCCGACATCACCGATGGGGAAGATCGGGCTCGGCATCTTCGGGCTCATGAGCG
5878	Db	CCTATATCGCCGACATCACCGATGGGGAAGATCGGGCTCGGCATCTTCGGGCTCATGAGCG	5937	Db	CCTATATCGCCGACATCACCGATGGGGAAGATCGGGCTCGGCATCTTCGGGCTCATGAGCG
4933	Qy	CTTTGTTTCGGCGTGGGTATGGTGGCAGGGCCCCCGTGGCCGGGGGACATGTTTGGGCGCCATCT	4992	Qy	CTTTGTTTCGGCGTGGGTATGGTGGCAGGGCCCCCGTGGCCGGGGGACATGTTTGGGCGCCATCT

Db 5938 CTTGTTTCGGCTGGGTATGTTGGCAGGCCCGTGGCGGGGACTGTTGGCGCACTCT 5997
QY CTTGTCATGCAC----- 5004
Db 5998 CTTTGGACCTGCAGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 6057
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 6058 CATTGCACAGATAAAAAATATATCATCATGAACAATAAAACTGTCTGCTTACATAAACAG 6117
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 6118 TAATACAGGGGTGTATGAGCCATATTCAACGGGAACGCTTTGCTCAAGSCCGGATT 6177
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 6178 AAATCCAAACATGGATGCTGATTTATATGGGTATAAAATGGGCTCGCGATAATGTGGGCA 6237
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 6238 ATCAGGTGCGACAACTCATTCGATTGTATGGGAAGCCCGATGCGCAGAGTTGTTTCTGAA 6297
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 6298 ACATGCAAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAACTGGCT 6357
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 6358 GACGGAATTTATGCCCTTTCCGACCATCAAGCATTTTATCCGTACTCCTGTATGATGATG 6417
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 6418 GTTACTCACCCTGCGATCCCGGGGAACACAGCATTCAGGTATTAGAGAATATCCTGA 6477
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 6478 TTCAGGTGAATAATTGTTGATGCGCTGCGAGTGTTCCTGCGCGTGTGCAATTCGATCC 6537
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 6538 TGTTGTAATTGCTCTTTTAAACAGCGATCGCGTATTTGCTGCTCAGGCGCAATCAG 6597
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 6598 AATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGCGCTGT 6657
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 6658 TGAACAAGTCTGGAAGAAATGATAGCTTTTGGCAATTCACCGGATTCAGTCGTAC 6717
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 6718 TCATGGTGATTTCTCACTTGATAACCTTATTTTTCAGAGGGGAATTAATAGTTGTAT 6777
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 6778 TGATGTTGACGAGTCGGAATCGCAGACGATACCGAGATCTTGCCATCTATGGAACGT 6837
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 6838 CCTCGTGATTTTCTCCTTCATTACAGAAACGGCTTTTCAAATAATGTTATGATAA 6897
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 6898 TCCTGATATGAATAAATTCAGTTTCATTGTGATGCTGATGATTTTCTAATCAGATT 6957
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 6958 GGTAAATTGGTTGTAACACTGGCAGAGCATTAAGTGAATTCAGCGGAGCGCGCTTGT 7017
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004

Db 7018 TGAATAAATCGAACTTTTGTGCTGAGTTGAAGATCAGATCAGCATCTTCCGACACGCA 7077
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7078 GACGTTCCGTGGCAAAGCAAAGTTCAAATCACCACCTGCTCCACCTACACAAAGCT 7137
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7138 CTATCAACCGTGGCTCCCTCACTTTCTGGCTGGATGATGGGGGATTCAGGCTGGTAT 7197
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7198 GAGTCAGCAACACCTTTTTCAGGAGCAGACCTCAGCGCCCCCCCCCTCGAGGTCCC 7257
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7258 ACAGCGCGGTGCTCAACGGCCTCAACCTACTACTGCGCTTCTTCTTAATGCAAGGATCG 7317
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7318 CATAAAGGAGAGCGTCAAGTATCTATGATTGGAAGTATGGAATGGTGTGATACCCGATTC 5132
Db 7319 CATAAAGGAGAGCGTCAAGTATCTATGATTGGAAGTATGGAATGGTGTGATACCCGATTC 7377
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7378 TTCACTCTCTTGAGGTCTCTATCAGATTATGCCCACTAAAGCAACCCGAGGAGGAGAT 5192
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7379 TTCACTCTCTTGAGGTCTCTATCAGATTATGCCCACTAAAGCAACCCGAGGAGGAGAT 7437
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7438 TTCACTCTCTTGAGGTCTCTATCAGATTATGCCCACTAAAGCAACCCGAGGAGGAGAT 7497
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7498 GTTATGACAGCAGAAATGTCTTCTTGGAGACAGTAAATGAAGTCCCAACCAATAAGAAA 5312
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7553 GTTATGACAGCAGAAATGTCTTCTTGGAGACAGTAAATGAAGTCCCAACCAATAAGAAA 7557
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7558 TCCTTGTATCAGAAACAACTTCTTGTTCGAACCTTTTCGGTGGCTTGAACATATAAA 5372
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7559 TCCTTGTATCAGAAACAACTTCTTGTTCGAACCTTTTCGGTGGCTTGAACATATAAA 7617
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7618 TGTAAGTGGATATGTCGGGTAGGAATGGAGCGGGCAAAATGCTTACCTTCGACCTTCA 5432
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7619 TGTAAGTGGATATGTCGGGTAGGAATGGAGCGGGCAAAATGCTTACCTTCGACCTTCA 7677
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7678 AGAGGTATGAGGTTTGTAGATCTGATCCAACTTCAGTGACAACTTCGTTGCTATTCGT 5492
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7679 TCAAAACCATTCGGAATCCAGAGAAATCAAAGTTGTTGCTACTATTGATCCAGGCCAGT 7737
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7738 TCAAAACCATTCGGAATCCAGAGAAATCAAAGTTGTTGCTACTATTGATCCAGGCCAGT 7797
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7798 GCGTCTTGAACTGACAAATAGTGTGCTGCTGTTTGGAGTCTCTTGTATGAATAAT 5552
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7799 GCGTCTTGAACTGACAAATAGTGTGCTGCTGTTTGGAGTCTCTTGTATGAATAAT 7797
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7857 GCGTCTTGAACTGACAAATAGTGTGCTGCTGTTTGGAGTCTCTTGTATGAATAAT 5612
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7858 CTAGTCTTGAATGACAAATAGTGTGCTGCTGTTTGGAGTCTCTTGTATGAATAAT 5672
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7859 CTAGTCTTGAATGACAAATAGTGTGCTGCTGTTTGGAGTCTCTTGTATGAATAAT 7917
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7918 CTTTAAACCGTTTAAAGGACAGTATGCTGCTGTTTAAACCCCAATCAGTCTGTA 5732
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7919 CTTTAAACCGTTTAAAGGACAGTATGCTGCTGTTTAAACCCCAATCAGTCTGTA 7977
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7978 GTCTGATCTCATCAACTGAGGGGCACTATCTTGTTTTGAAGAAATTTTCGGAGATGCG 5792
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7979 GTCTGATCTCATCAACTGAGGGGCACTATCTTGTTTTGAAGAAATTTTCGGAGATGCG 8037
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7980 ATATCGAGAAAGGATGAGTGTGTTTAAACGTTAAATTTTCTCAAGATCTTGCCTCG 5852
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7981 ATATCGAGAAAGGATGAGTGTGTTTAAACGTTAAATTTTCTCAAGATCTTGCCTCG 8097
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7982 CGCGTTTTCGGTGTAGCGGTGAAAACTCTCTGACACATGCGAGTCCCGGAGCGGTACAG 5912
QY CTTTGGACCTGCAGGGGGGGGGGGGAAGCACCAGTTGTGTCTCAAATCTCTGATGTA 5004
Db 7983 CGCGTTTTCGGTGTAGCGGTGAAAACTCTCTGACACATGCGAGTCCCGGAGCGGTACAG 8157

QY	5913	CTTGTCTCTAAGCGGATCCCGGAGCAGACAAGCCCGTCAGGCGCGCTCAGCGGCTGTG	5972	QY	6993	CTTCACTAGATCCCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATGA	7052
Db	8158	CTTGTCTCTAAGCGGATCCCGGAGCAGACAAGCCCGTCAGGCGCGCTCAGCGGCTGTG	8217	Db	9238	CTTCACTAGATCCCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATGA	9297
QY	5973	GGGGTGTGGGGCGACCATGACCCAGTACAGTAGCGATAGCGGAGTGTATCTGGCT	6032	QY	7053	GTAAACTTGTGTGACAGATTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTG	7112
Db	8218	GGGGTGTGGGGCGACCATGACCCAGTACAGTAGCGATAGCGGAGTGTATCTGGCT	8277	Db	9298	GTAAACTTGTGTGACAGATTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTG	9357
QY	6033	TAACTATCGGCATCAGAGCAGATTGTACTGAGAGTGACCATATGCGGTGTGAATACC	6092	QY	7113	TCTATTTCCTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGATAACTACGATACGGGA	7172
Db	8278	TAACTATCGGCATCAGAGCAGATTGTACTGAGAGTGACCATATGCGGTGTGAATACC	8337	Db	9358	TCTATTTCCTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGATAACTACGATACGGGA	9417
QY	6093	GCACAGATCGGTAAAGGAGAAATACCGCATCAGGCGCTCTTCGCTTCCTCGCTCACTGA	6152	QY	7173	GGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGAGACCCACGCTCACCGGCTCC	7232
Db	8338	GCACAGATCGGTAAAGGAGAAATACCGCATCAGGCGCTCTTCGCTTCCTCGCTCACTGA	8397	Db	9418	GGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGAGACCCACGCTCACCGGCTCC	9477
QY	6153	CTCGCTCGGCTCGGTCTGGCTCGCGGAGCGGTATCAGCTCACTCAAAAGCGGTAAAT	6212	QY	7233	AGATTTATCAGCAATAAAACCCAGCCGAGGAGGCGGAGCGCAGAAAGTGGTCTCTGCAAC	7292
Db	8398	CTCGCTCGGCTCGGTCTGGCTCGCGGAGCGGTATCAGCTCACTCAAAAGCGGTAAAT	8457	Db	9478	AGATTTATCAGCAATAAAACCCAGCCGAGGAGGCGGAGCGCAGAAAGTGGTCTCTGCAAC	9537
QY	6213	ACGCTATCCAGAAATCAGGGGTAACGCGAGGAAGAACATGACGCAAAAGGCCAGCA	6272	QY	7293	TTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGAGAGCTAGAGTAAGTAGTTCGCC	7352
Db	8458	ACGCTATCCAGAAATCAGGGGTAACGCGAGGAAGAACATGACGCAAAAGGCCAGCA	8517	Db	9538	TTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGAGAGCTAGAGTAAGTAGTTCGCC	9597
QY	6273	AAAGGCCAGGAACCGTAAAGGCGCGTGTGCTGCGCTTTTCCATAGGCTCCGCCCCC	6332	QY	7353	AGTTAATAGTTTGGCAACGTTTGGCAATTCCTGCGAGCATCGTGGTGTCAAGCTCGTC	7412
Db	8518	AAAGGCCAGGAACCGTAAAGGCGCGTGTGCTGCGCTTTTCCATAGGCTCCGCCCCC	8577	Db	9598	AGTTAATAGTTTGGCAACGTTTGGCAATTCCTGCGAGCATCGTGGTGTCAAGCTCGTC	9657
QY	6333	TGACGAGCATCAAAATTCGACGCTCAAGTCAGAGTGGGGAACCCGACGAGACTATA	6392	QY	7413	GTTTGTGTATGGCTTCATTTCAGCTCCGCTTCCCAACGATCAAGGCCAGTTACATGATCCCC	7472
Db	8578	TGACGAGCATCAAAATTCGACGCTCAAGTCAGAGTGGGGAACCCGACGAGACTATA	8637	Db	9658	GTTTGTGTATGGCTTCATTTCAGCTCCGCTTCCCAACGATCAAGGCCAGTTACATGATCCCC	9717
QY	6393	AAAGTACCAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTCTGTTCGCAACCTGCC	6452	QY	7473	CATGTTGTCAAAAAGCGGTTAGCTCTTCGGTCTCCGATCGTGTGTCAGAAAGTAGTT	7532
Db	8638	AAAGTACCAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTCTGTTCGCAACCTGCC	8697	Db	9718	CATGTTGTCAAAAAGCGGTTAGCTCTTCGGTCTCCGATCGTGTGTCAGAAAGTAGTT	9777
QY	6453	GCTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGCGCTTCTCTATGCTC	6512	QY	7533	GGCGGAGTGTATCACTCATCTGTTATGCGAGCATGTCATATTTCTTACTGTCTATGCC	7592
Db	8698	GCTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGCGCTTCTCTATGCTC	8757	Db	9778	GGCGGAGTGTATCACTCATCTGTTATGCGAGCATGTCATATTTCTTACTGTCTATGCC	9837
QY	6513	ACGCTAGGATATCAGTTCGGTGTAGTCTGCTCCAGCTGGGCTGTGTGACGA	6572	QY	7593	ATCCGTAAAGATGCTTTCTGTGATCTGTGATCTCAACCAAGTCACTCTGAGATAGTG	7652
Db	8758	ACGCTAGGATATCAGTTCGGTGTAGTCTGCTCCAGCTGGGCTGTGTGACGA	8817	Db	9838	ATCCGTAAAGATGCTTTCTGTGATCTGTGATCTCAACCAAGTCACTCTGAGATAGTG	9897
QY	6573	ACCCCGGTTACGCGGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAAACC	6632	QY	7653	TATGCGGCGACGAGTTGCTCTTGGCCGCGCTCAACACGGGATAATACCGCGCCACATAG	7712
Db	8818	ACCCCGGTTACGCGGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAAACC	8877	Db	9898	TATGCGGCGACGAGTTGCTCTTGGCCGCGCTCAACACGGGATAATACCGCGCCACATAG	9957
QY	6633	GGTAAGACAGCTTATCCGCTGCGACGACGCTGTTACAGGATAGCAGAGCGAG	6692	QY	7713	CAGAACTTTAAAGTCTCATCTGGAATAAGGTTCTTCGGGGCGAAAACTCTCAAGGAT	7772
Db	8878	GGTAAGACAGCTTATCCGCTGCGACGACGCTGTTACAGGATAGCAGAGCGAG	8937	Db	9958	CAGAACTTTAAAGTCTCATCTGGAATAAGGTTCTTCGGGGCGAAAACTCTCAAGGAT	10017
QY	6693	GTATGTAGCGGCTGTACAGAGTTCTTGAAGTGGTGGCTTAACTACGCGCTACCTAGAAG	6752	QY	7773	CTTACCGCTGTTGAGATCCAGTTCGATGTAAACCACTCGTGACCCAACTGATCTTCAGC	7832
Db	8938	GTATGTAGCGGCTGTACAGAGTTCTTGAAGTGGTGGCTTAACTACGCGCTACCTAGAAG	8997	Db	10018	CTTACCGCTGTTGAGATCCAGTTCGATGTAAACCACTCGTGACCCAACTGATCTTCAGC	10077
QY	6753	GACAGATTGTTGATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAG	6812	QY	7833	ATCTTTTACTTTTCCAGCGGTTTCTGGGTGAGCAAAAAAGGAGGCAAAATGCCGCAAA	7892
Db	8998	GACAGATTGTTGATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAG	9057	Db	10078	ATCTTTTACTTTTCCAGCGGTTTCTGGGTGAGCAAAAAAGGAGGCAAAATGCCGCAAA	10137
QY	6813	CTCTTCATCCGCAAAACCAACCAACCGCTGAGCGGTGGTGGTGGTGGTGGTGGTGGT	6872	QY	7893	AAAGGGAATGAGGCGGACGAGGAAATGTTGAATCTCATCTCTCTCTCTCTCTCTCTCT	7952
Db	9058	CTCTTCATCCGCAAAACCAACCAACCGCTGAGCGGTGGTGGTGGTGGTGGTGGTGGT	9117	Db	10138	AAAGGGAATGAGGCGGACGAGGAAATGTTGAATCTCATCTCTCTCTCTCTCTCTCTCT	10197
QY	6873	GATTACCGCAGAAAAAGGATCTCAAGAGATCTCTTTGATCTCTTTCTACGGGTCTGA	6932	QY	7953	TTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTAGAA	8012
Db	9118	GATTACCGCAGAAAAAGGATCTCAAGAGATCTCTTTGATCTCTTTCTACGGGTCTGA	9177	Db	10198	TTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTAGAA	10257
QY	6933	CGCTCAGTGGAGCAAACTCAGTTTAAAGGATTTTGGTCATGAGATTATCAAAAGGAT	6992	QY	8013	AAATAAACAATAGGGGTTCCGGGCACTTTCCCGAAAAAGTCCACCTGAGCTCTAAGA	8072
Db	9178	CGCTCAGTGGAGCAAACTCAGTTTAAAGGATTTTGGTCATGAGATTATCAAAAGGAT	9237	Db	10258	AAATAAACAATAGGGGTTCCGGGCACTTTCCCGAAAAAGTCCACCTGAGCTCTAAGA	10317
				QY	8073	AACCATTTATCATGACATTAACCTATAAAATAGGCGTATCAGGAGGCCCTTTCTGCTCT	8132

Db	10318	AACCAATTATTATCATGACATTAACTATAAAAAATAGGCGTATCAGAGGCCCTTTCGTCT	10377
Qy	8133	TCAGAATAATTCTCATGTTTGACAGCTTATCATCGATAAGCTGACTCATGTTGGTATT	8192
Db	10378	TCAGAATAATTCTCATGTTTGACAGCTTATCATCGATAAGCTGACTCATGTTGGTATT	10437
Qy	8193	GTGAAATAGACGACATCGGAACACTGAAAAATAACAGTTATTATTTCG	8241
Db	10438	GTGAAATAGACGACATCGGAACACTGAAAAATAACAGTTATTATTTCG	10486

Search completed: September 24, 2004, 04:29:51
Job time : 21002.8 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 04:29:55 ; Search time 1464.26 Seconds
(without alignments)
17786.465 Million cell updates/sec

Title: US-10-030-390-1
Perfect score: 5142
Sequence: 1 gaattcgagtcgtaccg.....ccgtctgtaagtccttaag 5142

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3337386 seqs, 2532474682 residues

Total number of hits satisfying chosen parameters: 6674772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:**

- 1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq*
- 2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq*
- 3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq*
- 4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq*
- 5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq*
- 6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq*
- 7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq*
- 8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq*
- 9: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq*
- 10: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq*
- 11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq*
- 12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq*
- 13: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq*
- 14: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq*
- 15: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq*
- 16: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq*
- 17: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq*
- 18: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq*
- 19: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4232.8	82.3	10929	15	US-10-032-393-7
2	1493	29.0	5865	13	US-10-402-466A-2
3	1493	29.0	6894	13	US-10-402-466A-1
4	1493	29.0	7336	13	US-10-462-128-3
5	1493	29.0	7336	15	US-10-084-814-3
6	1493	29.0	7336	16	US-10-442-013-4
7	1493	29.0	8198	13	US-10-402-466A-7
8	1493	29.0	9286	13	US-10-402-466A-10
9	1487.6	28.9	5302	16	US-10-385-415-2
10	1487.6	28.9	5767	16	US-10-385-415-4
11	1207.6	23.5	5793	10	US-09-869-855A-3
12	1207.6	23.5	5943	10	US-09-869-855A-2
13	981.2	19.1	6852	15	US-10-032-393-16
14	981.2	19.1	6852	15	US-10-032-393-53

C 15	978.4	19.0	3867	13	US-10-466-642-24	Sequence 24, Appl
C 16	978.4	19.0	3867	17	US-10-466-717-24	Sequence 24, Appl
C 17	978	19.0	6725	13	US-10-466-642-22	Sequence 22, Appl
C 18	978	19.0	6725	17	US-10-466-717-22	Sequence 22, Appl
C 19	978	19.0	6806	13	US-10-466-642-23	Sequence 23, Appl
C 20	978	19.0	6806	17	US-10-466-717-23	Sequence 23, Appl
C 21	978	19.0	6886	13	US-10-466-642-20	Sequence 20, Appl
C 22	978	19.0	6886	17	US-10-466-717-20	Sequence 20, Appl
C 23	978	19.0	7140	13	US-10-466-642-21	Sequence 21, Appl
C 24	978	19.0	7140	17	US-10-466-717-21	Sequence 21, Appl
C 25	978	19.0	8320	9	US-09-813-453A-71	Sequence 71, Appl
C 26	978	19.0	8803	17	US-10-466-717-27	Sequence 27, Appl
C 27	977	19.0	7701	17	US-10-466-641-25	Sequence 25, Appl
C 28	977	19.0	7926	17	US-10-466-641-24	Sequence 24, Appl
C 29	943	18.3	10186	15	US-10-234-329-3	Sequence 3, Appl
C 30	867	16.9	3156	13	US-10-681-086-1	Sequence 1, Appl
C 31	866.8	16.9	2517	9	US-09-893-600-1	Sequence 1, Appl
C 32	866.8	16.9	2517	9	US-09-893-499-1	Sequence 1, Appl
C 33	866.8	16.9	2517	9	US-09-894-030-1	Sequence 1, Appl
C 34	771	15.0	6837	10	US-09-928-847B-49	Sequence 49, Appl
C 35	759	14.8	759	17	US-10-718-628-10	Sequence 10, Appl
C 36	692	13.5	692	9	US-09-974-300-6910	Sequence 6910, Ap
C 37	541	10.5	6405	10	US-09-869-855A-1	Sequence 1, Appl
C 38	347.2	6.8	8704	13	US-10-466-642-25	Sequence 25, Appl
C 39	347.2	6.8	8704	17	US-10-466-717-25	Sequence 25, Appl
C 40	347.2	6.8	9292	9	US-09-813-453A-74	Sequence 74, Appl
C 41	347.2	6.8	9396	9	US-09-813-453A-73	Sequence 73, Appl
C 42	198.6	3.9	447	16	US-10-166-349-11	Sequence 11, Appl
C 43	198.6	3.9	447	16	US-10-166-349-19	Sequence 19, Appl
C 44	170.2	3.3	4914	16	US-10-310-740-11	Sequence 11, Appl
C 45	162	3.2	5502	9	US-09-736-457-785	Sequence 785, App

ALIGNMENTS

RESULT 1

US-10-032-393-7
; Sequence 7, Application US/10032393
; Publication No. US20030027286A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Wall, Daniel
; APPLICANT: Gross, Molly
; TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
; FILE REFERENCE: ELITRA.010A
; CURRENT APPLICATION NUMBER: US/10/032,393
; CURRENT FILING DATE: 2001-12-21
; PRIOR FILING DATE: 2000-12-27
; PRIOR FILING DATE: 2000-12-27
; PRIOR FILING DATE: 2000-12-27
; PRIOR FILING DATE: 2001-09-06
; PRIOR FILING DATE: 2001-09-06
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 10929
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Vector pep25
US-10-032-393-7

Query Match 82.3%; Score 4232.8; DB 15; Length 10929;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 4323; Conservative 0; Mismatches 17; Indels 10; Gaps 7;

Qy 651 CTGATTTTCATTTTTCGATTTCTACAACTGCATATGTAATCGCTCTCTTTTA 710
Db 6491 CTGATTTTCATTTTTCGATTTCTACAACTGCATATGTAATCGCTCTCTTTTA 6550
Qy 711 GGTGGCAAAATGTGAGGCATTTTCGCTCTTTTCGGCGAGGCTAGTTACCCCTTAAGTTAT 770

|||||
6551 GGTGCAAAATGTGAGGCAATTTTCGCTCTTTCGGCGAGGCTAGTTACCCCTTAAGTTAT 6610
QY
771 TGGTATGACTGGTTTAAAGCCCAAAAAGTTGCTTTTTCGTACCTATTAATGATCGTT 830
Db
6611 TGGTATGACTGGTTTAAAGCCCAAAAAGTTGCTTTTTCGTACCTATTAATGATCG-- 6668
QY
831 TTAGAAAACCGACTGTAAAAGTACAGTCGGCATTTCTCATATTAATAAAGCCAGTCAT 890
Db
6669 TTAGAAAACCGACTGTAAAAGTACAGTCGGCATTTCTCATATTAATAAAGCCAGTCAT 6728
QY
891 TAGGCTATCTGACAAATTCCTGAATAGAGTTTATAAACAATTCCTGCGATGATAACCATCAC 950
Db
6729 TAGGCTATCTGACAAATTCCTGAATAGAGTTTATAAACAATTCCTGCGATGATAACCATCAC 6788
QY
951 AAAAGAAATGATGTACCTGTAAAAGTACAGTCGGCTAAATATATTAATTAATTAATGA 1010
Db
6789 AAAAGAAATGATGTACCTGTAAAAGTACAGTCGGCTAAATATATTAATTAATTAATGA 6848
QY
1011 ATTTTCCTGCTGTAATAATGGTGAAGTAATTAATTAATTAATTAATTAATTAATTAATTA 1070
Db
6849 ATTTTCCTGCTGTAATAATGGTGAAGTAATTAATTAATTAATTAATTAATTAATTAATTA 6908
QY
1071 ACCCAGTAATGAAGTCATGGAATAATAGAAAGAGAAAAGCAATTTTCAGGTATAGGTG 1130
Db
6909 ACCCAGTAATGAAGTCATGGAATAATAGAAAGAGAAAAGCAATTTTCAGGTATAGGTG 6968
QY
1131 TTTTCGGGAAAACAATTTCCCGAACCATTTATTTCTCTACATCAGAAAGGTATATAATCAT 1190
Db
6969 TTTTCGGGAAAACAATTTCCCGAACCATTTATTTCTCTACATCAGAAAGGTATATAATCAT 7028
QY
1191 AAAACTCTTTGAAGTCATTTCTTACAGGAGTCGAAATACAGAGAAATGTTTTAGATACAC 1250
Db
7029 AAAACTCTTTGAAGTCATTTCTTACAGGAGTCGAAATACAGAGAAATGTTTTAGATACAC 7088
QY
1251 CATCAAAAATTTGATAAAGTGGCTCTAACTTTACCAATACCTAACTCCGCGGTAT 1310
Db
7089 CATCAAAAATTTGATAAAGTGGCTCTAACTTTACCAATACCTAACTCCGCGGTAT 7148
QY
1311 TGTAAACCAAGTCTTAAAGCTGATTTTGAAGTTTATCACCCCTTGTCACTAAGAAATAAATG 1370
Db
7149 TGTAAACCAAGTCTTAAAGCTGATTTTGAAGTTTATCACCCCTTGTCACTAAGAAATAAATG 7208
QY
1371 CAGGCTAAAATTTATATCTCTTGTGTTTATGTTTCGTTATGATTAACCACTAATCAATTT 1430
Db
7209 CAGGCTAAAATTTATATCTCTTGTGTTTATGTTTCGTTATGATTAACCACTAATCAATTT 7268
QY
1431 CTGTGTTTATCTAAAGTCTGTTTGTGTTTCAAAATATGATTAATATCTCTTTTCTCT 1490
Db
7269 CTGTGTTTATCTAAAGTCTGTTTGTGTTTCAAAATATGATTAATATCTCTTTTCTCT 7328
QY
1491 TCCAAATGCTTAAATCAATTTTATTAAGTTTCATTTGATATGCTTCTTAATTTTATCT 1550
Db
7329 TCCAAATGCTTAAATCAATTTTATTAAGTTTCATTTGATATGCTTCTTAATTTTATCT 7388
QY
1551 AAAGTGAATTTAGGAGGCTTACTTGTCTCTTCTTCAATAGAAATCAATCTTTTAAA 1610
Db
7389 AAAGTGAATTTAGGAGGCTTACTTGTCTCTTCTTCAATAGAAATCAATCTTTTAAA 7448
QY
1611 AGTCAATATTAAGTGAACATAAATATATATTTTAAATAATATCCCATTTTATCCAAATTTTC 1670
Db
7449 AGTCAATATTAAGTGAACATAAATATATATTTTAAATAATATCCCATTTTATCCAAATTTTC 7508
QY
1671 GTTTGTTGAATAGGAGGCTTGTAGTTGAAGAT-AAAAGACCAATTAATAAATGTTGTT 1729
Db
7509 GTTTGTTGAATAGGAGGCTTGTAGTTGAAGATTAATAAAGACCAATTAATAAATGTTGTT 7568
QY
1730 CTTTGTGTTTAAAGGATTTGAGGCTAGGAGGATTTTCTTTTCTTTCTTATCTTGA 1789
Db
7569 CTTTGTGTTTAAAGGATTTGAGGCTAGGAGGATTTTCTTTTCTTTCTTATCTTGA 7628
QY
1790 TAATAAGGTAATTTTCGGGATAGACTGTAAATCTTCACGATAAATAATCCCTTTTC 1849

Db
7629 TAATAAGGTAATTTTCGGGATAGACTGTAAATCTTCACGATAAATAATCCCTTTTC 7688
QY
1850 ATTTTCTAATGTAATCTATTAACCTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 1909
Db
7689 ATTTTCTAATGTAATCTATTAACCTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 7748
QY
1910 TTTCTTATTAACGAAAATGGCCGATTTAAGCACACCCCTTTATTCGGTTAATTCGCCAATGA 1969
Db
7749 TTTCTTATTAACGAAAATGGCCGATTTAAGCACACCCCTTTATTCGGTTAATTCGCCAATGA 7808
QY
1970 CAGCCATGATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2029
Db
7809 CAGCCATGATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7868
QY
2030 AATTTATTAAGAGTCTATCGTTTCAAAATGGTATGTCGGTTTGAACATCCCACTATATATCCGT 2089
Db
7869 AATTTATTAAGAGTCTATCGTTTCAAAATGGTATGTCGGTTTGAACATCCCACTATATATCCGT 7928
QY
2090 GTCGTTCTGTCCACTCTCGAATCCCAATTCAGAAATTTCTAGCGATTTCCAGAAAGTTTCT 2149
Db
7929 GTCGTTCTGTCCACTCTCGAATTCAGAAATTTCTAGCGATTTCCAGAAAGTTTCT 7988
QY
2150 CAGAGTCGAAAAGTTGACCGAGCATTTACGAACTGSCACAGATGTCATACCTGAAGGAA 2209
Db
7989 CAGAGTCGAAAAGTTGACCGAGCATTTACGAACTGSCACAGATGTCATACCTGAAGGAA 8048
QY
2210 GATCTGATTTGCTTAACTGCTTCAAGTAAAGCCGAGCGCTCTCGTATATAACAGATGCGAT 2269
Db
8049 GATCTGATTTGCTTAACTGCTTCAAGTAAAGCCGAGCGCTCTCGTATATAACAGATGCGAT 8108
QY
2270 GATCGAGACCAATCAACATGCGACCTGCGATTTGCTTACCTGTACAGTCAAGATGCTAGAA 2329
Db
8109 GATCGAGACCAATCAACATGCGACCTGCGATTTGCTTACCTGTACAGTCAAGATGCTAGAA 8168
QY
2330 ATGTTGTCTGCTTGTGCAACGAAATATTACGCAATTTGCTTGCATATTTCAAAACAGCTCTT 2389
Db
8169 ATGTTGTCTGCTTGTGCAACGAAATATTACGCAATTTGCTTGCATATTTCAAAACAGCTCTT 8228
QY
2390 CTACGATAAGGCGCAAAATCGCATCTGTGAGACGTTTGGGCTTCTACCGATTTAGCAGTTT 2449
Db
8229 CTACGATAAGGCGCAAAATCGCATCTGTGAGACGTTTGGGCTTCTACCGATTTAGCAGTTT 8288
QY
2450 GATACACTTTTCTTAAGTATCCACTGCAATCATATAATCGGAAAATAGAGAAAATTTGAC 2509
Db
8289 GATACACTTTTCTTAAGTATCCACTGCAATCATATAATCGGAAAATAGAGAAAATTTGAC 8348
QY
2510 CATGTGTAAGCGGCCAAATCTGATTTCCACTGAGATGCAATAATCTAGTAGAATCTCTTCCG 2569
Db
8349 CATGTGTAAGCGGCCAAATCTGATTTCCACTGAGATGCAATAATCTAGTAGAATCTCTTCCG 8408
QY
2570 TATCAAAATTTCTTCCACTTCCACTCAGCGTTTGTCCATTTATGCTGCTGAATCTGCTT 2629
Db
8409 TATCAAAATTTCTTCCACTTCCACTCAGCGTTTGTCCATTTATGCTGCTGAATCTGCTT 8468
QY
2630 CCTCTGTTGACATGACACACATCATCTCAATATCCGAAATAGGCGCCCATCAGTCTGACGAC 2689
Db
8469 CCTCTGTTGACATGACACACATCATCTCAATATCCGAAATAGGCGCCCATCAGTCTGACGAC 8528
QY
2690 CAGAGAGCCATAAACCACCAATAGCTTTAAACATCATCCCATTTATTTATCCATATTCGTT 2749
Db
8529 CAGAGAGCCATAAACCACCAATAGCTTTAAACATCATCCCATTTATTTATCCATATTCGTT 8588
QY
2750 CTTAATTTTATGAACAATCTTTCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2809
Db
8589 CTTAATTTTATGAACAATCTTTCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 8648
QY
2810 CTTAATTTTATGAACAATTTTCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2869
Db
8649 CTTAATTTTATGAACAATTTTCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 8708
QY
2870 GAGCAGCGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2929
Db
8709 GAGCAGCGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 8768

GENERAL INFORMATION:
APPLICANT: Park, Sukjoon
APPLICANT: Giri, Lallan
TITLE OF INVENTION: RECOMBINANT IMMUNOGENIC COMPOSITIONS AND METHODS FOR PROTECTING AGAINST LETHAL INFECTIONS FROM BACILLUS ANTHRACIS
FILE REFERENCE: 18933-00005
CURRENT APPLICATION NUMBER: US/10/402,466A
CURRENT FILING DATE: 2003-03-20
PRIOR APPLICATION NUMBER: 60/372,152
PRIOR FILING DATE: 2002-04-12
NUMBER OF SEQ ID NOS: 34
SEQ ID NO 2
LENGTH: 5865
TYPE: DNA
ORGANISM: Artificial
FEATURE:
NAME/KEY:
LOCATION:
OTHER INFORMATION: Artificial DNA sequence to be used as one of the two backbones of E. coli-Bacillus anthracis shuttle vectors. Designated as pBP II
US-10-402-466A-2

Query Match 29.0%; Score 1493; DB 13; Length 5865;
Best Local Similarity 98.1%; Pred. No. 2.7e-282;
Matches 1511; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY	1778	TTCTTATCTTGATAATAAGGTTAACTATTGCGGGATAGACTGTAAACATTTCTCAGCGATA	1837
DB	5526	TACTCTTTTAAATATCCCGGACTGCGAATGCGGGATAGACTGTAAACATTTCTCAGCGATA	5467
QY	1838	AAATCCCTTTTCAATTTCTTAATGTAATCTAATACCTTATTATTAAATTCAGTTCAT	1897
DB	5466	AAATCCCTTTTCAATTTCTTAATGTAATCTAATACCTTATTATTAAATTCAGTTCAT	5407
QY	1898	AATTAATCTTTTCTTATTAGGAAATGCGCGATTAAGCACACCCCTTTATTCGGTT	1957
DB	5406	AATTAATCTTTTCTTATTAGGAAATGCGCGATTAAGCACACCCCTTTATTCGGTT	5347
QY	1958	AATGGCCATGACGCCATGATAATTAATCTAATCTAGGAGAAATTAATAATAGTAACC	2017
DB	5346	AATGGCCATGACGCCATGATAATTAATCTAATCTAGGAGAAATTAATAATAGTAACC	5287
QY	2018	AACATGATTAAACAATATTAGAGTTCATCGTTCAAAATGGTATGGTTTGACACATCCA	2077
DB	5286	AACATGATTAAACAATATTAGAGTTCATCGTTCAAAATGGTATGGTTTGACACATCCA	5227
QY	2078	CTATATATCCGTGCTGTTCTGTCCTCACTCTGAATCCCATTCAGAAATTCCTAGCGATT	2137
DB	5226	CTATATATCCGTGCTGTTCTGTCCTCACTCTGAATCCCATTCAGAAATTCCTAGCGATT	5167
QY	2138	CCAGAAGTTTCTCAGATCGGAAATGACACAGACATTAAGAACTGGCAGACAGATGGTCA	2197
DB	5166	CCAGAAGTTTCTCAGATCGGAAATGACACAGACATTAAGAACTGGCAGACAGATGGTCA	5107
QY	2198	AACTGAGAGAGATCTGATGCTTAACTGCTTCAAGTTAAGCAGGCGCTGCTGCTAT	2257
DB	5106	AACTGAGAGAGATCTGATGCTTAACTGCTTCAAGTTAAGCAGGCGCTGCTGCTAT	5047
QY	2258	AACAGATGCGATGATGACAGACCAATCAACATGGCACTGCCATTCCTACCTGACATCA	2317
DB	5046	AACAGATGCGATGATGACAGACCAATCAACATGGCACTGCCATTCCTACCTGACATCA	4987
QY	2318	AGGATGCTAGAAATCTGTCGGTCTTCACACGATTAATAGCCATTTGCTGCTGATTT	2377
DB	4986	AGGATGCTAGAAATCTGTCGGTCTTCACACGATTAATAGCCATTTGCTGCTGATTT	4927
QY	2378	CAACAGCTCTTCTACGATAGGGCAGAAATCGCATCGTGAAGCTTTGGGCTTCTACCG	2437
DB	4926	CAACAGCTCTTCTACGATAGGGCAGAAATCGCATCGTGAAGCTTTGGGCTTCTACCG	4867
QY	2438	ATTAGAGTGTGATACACTTCTCTAGATATCAACCTGATCATTAATCGGCAAAATAG	2497
DB	4866	ATTAGAGTGTGATACACTTCTCTAGATATCAACCTGATCATTAATCGGCAAAATAG	4807

QY	2498	AGAAAAATTGACATGTGTAAAGCGGCAATCTGATTCCACCTGAGATCGATAATCTAGTA	2557
DB	4806	AGAAAAATTGACATGTGTAAAGCGGCAATCTGATTCCACCTGAGATCGATAATCTAGTA	4747
QY	2558	GAATCTCTTCGTATCAAAATTCACCTTCCACCTTCCACCTTCCACCTTCCACCTTCCACCT	2617
DB	4746	GAATCTCTTCGTATCAAAATTCACCTTCCACCTTCCACCTTCCACCTTCCACCTTCCACCT	4687
QY	2618	TGAATCTCTTCGTATCAAAATTCACCTTCCACCTTCCACCTTCCACCTTCCACCTTCCACCT	2677
DB	4686	TGAATCTCTTCGTATCAAAATTCACCTTCCACCTTCCACCTTCCACCTTCCACCTTCCACCT	4627
QY	2678	CAGTCTGACGACCAAGAGAGCCATAAACAACCAATAGCTTAAACATCATCCCCCATATTTAT	2737
DB	4626	CAGTCTGACGACCAAGAGAGCCATAAACAACCAATAGCTTAAACATCATCCCCCATATTTAT	4567
QY	2738	CCAATATTCGTTCCCTTAATTTCAATGAAACAAATCTTCAATCTTCTCTCTAGTCATATTA	2797
DB	4566	CCAATATTCGTTCCCTTAATTTCAATGAAACAAATCTTCAATCTTCTCTCTAGTCATATTA	4507
QY	2798	TTGGTCCATTCATCTCTCAATCCCTTTTCAAGATAATTTAGATTTGCTTTCTTAAATA	2857
DB	4506	TTGGTCCATTCATCTCTCAATCCCTTTTCAAGATAATTTAGATTTGCTTTCTTAAATA	4447
QY	2858	AGAAATATTTGGAGAGACCGTTTCTTATTCAAGCTATTAAATACTCGTCTTCTTAAGCATCC	2917
DB	4446	AGAAATATTTGGAGAGACCGTTTCTTATTCAAGCTATTAAATACTCGTCTTCTTAAGCATCC	4387
QY	2918	TTCAATCCCTTTTAAATAACCAATTTAGCATCTTAATCTTCAACAACTGCGGTTTGTGA	2977
DB	4386	TTCAATCCCTTTTAAATAACCAATTTAGCATCTTAATCTTCAACAACTGCGGTTTGTGA	4327
QY	2978	ACTACTCTTTTAAATAAATTTTCCCGTCCCAATTCACATTCGCAATTAAGAAATC	3037
DB	4326	ACTACTCTTTTAAATAAATTTTCCCGTCCCAATTCACATTCGCAATTAAGAAATC	4267
QY	3038	CATCTTCATCGGCTTTTCGTCAATCATCTGTATGAATCAAAATCGCCTTCTCTGTGTAT	3097
DB	4266	CATCTTCATCGGCTTTTCGTCAATCATCTGTATGAATCAAAATCGCCTTCTCTGTGTAT	4207
QY	3098	CAAGGTTTAAATTTTATGATTTCTTTTAAACAAACCACTAGGAGATTAACTTTTAC	3157
DB	4206	CAAGGTTTAAATTTTATGATTTCTTTTAAACAAACCACTAGGAGATTAACTTTTAC	4147
QY	3158	GGTAAACCTTCTCCAAATCAGACAAACGTTTCAAAATCTTCTTCAATCATCGGTCA	3217
DB	4146	GGTAAACCTTCTCCAAATCAGACAAACGTTTCAAAATCTTCTTCAATCATCGGTCA	4087
QY	3218	TAAATCCGATCTTCAAGATATTTTGCAGTTTTCGTCATTCGCGATTGTATATCGG	3277
DB	4086	TAAATCCGATCTTCAAGATATTTTGCAGTTTTCGTCATTCGCGATTGTATATCGG	4027
QY	3278	ATTATATTTATTTTTCGGTATTTTATTTTAAACGTCCTCA	3318
DB	4026	ATTATATTTATTTTTCGGTATTTTATTTTAAACGTCCTCA	3986

RESULT 3
US-10-402-466A-1/c
Sequence 1, Application US/10402466A
Publication No. US20040028695A1
GENERAL INFORMATION:
APPLICANT: Park, Sukjoon
APPLICANT: Giri, Lallan
TITLE OF INVENTION: RECOMBINANT IMMUNOGENIC COMPOSITIONS AND METHODS FOR PROTECTING AGAINST LETHAL INFECTIONS FROM BACILLUS ANTHRACIS
FILE REFERENCE: 18933-00005
CURRENT APPLICATION NUMBER: US/10/402,466A
CURRENT FILING DATE: 2003-03-20
PRIOR APPLICATION NUMBER: 60/372,152
PRIOR FILING DATE: 2002-04-12
NUMBER OF SEQ ID NOS: 34

```

; SEQ ID NO 1
; LENGTH: 6694
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; NAME/KEY:
; LOCATION:
; OTHER INFORMATION: Artificial DNA sequence to be used as one of the two backbones of
; OTHER INFORMATION: E. coli-Bacillus anthracis shuttle vectors. Designated as pBP 1.
US-10-402-466A-1

```

Query Match	29.0%;	Score 1493;	DB 13;	Length 6694;
Best Local Similarity	98.1%;	Pred. No. 2.9e-282;		
Matches 1511;	Conservative	0;	Mismatches 30;	Indels 0;
				Gaps 0;

Qy	1778	TTCTTATCTTCGATGAATAAAGGGTAACATATTGCCGGGATAGACTGTAAACATTTCTCACGCATA	1831
Db	6355	TACTCTTTTTTAATATCCCGAGCTGGCAATGCCGGGATAGACTGTAAACATTTCTCACGCATA	6396
Qy	1838	AAATCCCGCTTCATCTTTTCTAATGTAAATCTATTACCTCTATTATTAAATTCAAATTCGCTCAT	1897
Db	6295	AAATCCCGCTTCATCTTTTCTAATGTAAATCTATTACCTCTATTATTAAATTCAAATTCGCTCAT	6236
Qy	1898	AAATTAATCCCTTTTCTTATTACGCAGAAATAGCCCGGATTTAAGCACACCCCTTTATTCGGTT	1957
Db	6235	AAATTAATCCCTTTTCTTATTACGCAGAAATAGCCCGGATTTAAGCACACCCCTTTATTCGGTT	6176
Qy	1958	AATCGCGCATGACAGCCGATGATTAATCTAATATCTAGGAGAAGTTAATAATACGTAAACC	2017
Db	6175	AATCGCGCATGACAGCCGATGATTAATCTAATATCTAGGAGAAGTTAATAATACGTAAACC	6116
Qy	2018	AACATGATTAAACAATATTAGAGTCTATCGTTCAAAATGGTATGCGTTTGTGACACATCCA	2077
Db	6115	AACATGATTAAACAATATTAGAGTCTATCGTTCAAAATGGTATGCGTTTGTGACACATCCA	6056
Qy	2078	CTATATATCCGTTGCTGTCCTCCACTCCTCGAATCCCATTCGCAAAATTTCTCTAGCGCAT	2137
Db	6055	CTATATATCCGTTGCTGTCCTCCACTCCTCGAATCCCATTCGCAAAATTTCTCTAGCGCAT	5996
Qy	2138	CCAGAAGTTTCTCAGAGTCGGAAGAAGTTGACACAGACATTTACGAACTGGCAGACATGTCAT	2197
Db	5995	CCAGAAGTTTCTCAGAGTCGGAAGAAGTTGACACAGACATTTACGAACTGGCAGACATGTCAT	5936
Qy	2198	AACCTGAAGGAAGATCTGATTGCTTTAACTGCTTCAGTTTAAGACCGAAGCGCTCGTCGAT	2257
Db	5935	AACCTGAAGGAAGATCTGATTGCTTTAACTGCTTCAGTTTAAGACCGAAGCGCTCGTCGAT	5876
Qy	2258	AACAGATCGGATGATGCAGAGCAATCAACATGGCACCTCGCCATTCGTACCTGTACAGTCA	2317
Db	5875	AACAGATCGGATGATGCAGAGCAATCAACATGGCACCTCGCCATTCGTACCTGTACAGTCA	5816
Qy	2318	AGGATGTPAGAAATGTTGCTCGGTCTTGCAACAGAAATTAAGCCATTTGCTCGATATT	2377
Db	5815	AGGATGTPAGAAATGTTGCTCGGTCTTGCAACAGAAATTAAGCCATTTGCTCGATATT	5756
Qy	2378	CACACAGCTCTCTACGATAGGGCACAAATCGCATCGTGAAGAGTTGGGCTTCTACCG	2437
Db	5755	CACACAGCTCTCTACGATAGGGCACAAATCGCATCGTGAAGAGTTGGGCTTCTACCG	5696
Qy	2438	ATTTAGCAGTTTGTATACACTTTTCTCTAAGTATCCAACCTGAAATCATAAATCGGCAAAATAG	2497
Db	5695	ATTTAGCAGTTTGTATACACTTTTCTCTAAGTATCCAACCTGAAATCATAAATCGGCAAAATAG	5636
Qy	2498	AGAAAAATGACCATGTGTAGCGGCCAATCTGATTCACCTGAGATGCATAATCTAGTA	2557
Db	5635	AGAAAAATGACCATGTGTAGCGGCCAATCTGATTCACCTGAGATGCATAATCTAGTA	5576
Qy	2558	GAAATCTCTCGGTATCAAAATTCACCTCCACTCCACCGGTTGTCATTCATGGC	2617
Db	5575	GAAATCTCTCGGTATCAAAATTCACCTCCACTCCACCGGTTGTCATTCATGGC	5516
Qy	2618	TGAATCTCGTTCCTCTGTTGATGACACATCATCTCAATATCCGAATAGGCCCAT	2677

Db	5515	TGA	ACTCTGCTT	CCTCTGTTG	CATGACAC	CATCATCT	CAATAT	CGGAATAGG	CCCAT	5445
Qy	2678	CAG	CTGAGCAG	CCAGAGAG	CCCAATAAC	CCAAATAG	CGCTTAA	CATCAT	CCCCATATTAT	2737
Db	5455	CAG	CTCAGC	CAAGAGAG	CCCAATAAC	CCAAATAG	CGCTTAA	CATCAT	CCCCATATTAT	5396
Qy	2738	CCA	ATATTCGT	TTCCTTAAT	TTCATGAAC	CAATCTT	CTTCTCT	CTAGTCAT	TATTA	2797
Db	5395	CCA	ATATTCGT	TTCCTTAAT	TTCATGAAC	CAATCTT	CTTCTCT	CTAGTCAT	TATTA	5336
Qy	2798	T	TGGTCAT	TACTAT	TCTCAT	TCCCTTTT	TCAGAT	TAATTT	TAGATTTGCTTTCTTAATA	2857
Db	5335	T	TGGTCAT	TACTAT	TCTCAT	TCCCTTTT	TCAGAT	TAATTT	TAGATTTGCTTTCTTAATA	5276
Qy	2858	AGA	ATATTTGG	AGACAC	CGCTTCTTA	TTCAGCTATT	TAATAAC	TCTGCTTCT	CTAAGCATCC	2917
Db	5275	AGA	ATATTTGG	AGACAC	CGCTTCTTA	TTCAGCTATT	TAATAAC	TCTGCTTCT	CTAAGCATCC	5216
Qy	2918	TT	CAATCCTTTT	TATAACAA	TATATAG	CATCTAAT	CTTCAACAA	CTGCGCGCTT	TGTTGA	2977
Db	5215	TT	CAATCCTTTT	TATAACAA	TATATAG	CATCTAAT	CTTCAACAA	CTGCGCGCTT	TGTTGA	5156
Qy	2978	A	CTACTCTTT	TAAATAA	TAATTTTCG	TTCGCCAAT	TCCCAT	TGCAAT	TGCAATATAGAAATC	3037
Db	5155	A	CTACTCTTT	TAAATAA	TAATTTTCG	TTCGCCAAT	TCCCAT	TGCAAT	TGCAATATAGAAATC	5096
Qy	3038	CAT	CTTCAT	CGCGCTTTT	TCGTCA	TCACTGTAT	GAATCAA	AATGCCCTTCT	TCTGTGTCAT	3097
Db	5095	CAT	CTTCAT	CGCGCTTTT	TCGTCA	TCACTGTAT	GAATCAA	AATGCCCTTCT	TCTGTGTCAT	5036
Qy	3098	CA	AGGTTAA	TTTTTAT	TATATTTCTTTT	TAAACAC	CACCAT	TAGGAGAT	TAACCTTTTAC	3157
Db	5035	CA	AGGTTAA	TTTTTAT	TATATTTCTTTT	TAAACAC	CACCAT	TAGGAGAT	TAACCTTTTAC	4976
Qy	3158	GG	GTAAAC	CCCTTCT	CCAAAT	CAGACAA	ACGTTTCAA	AATCTTTTCTTCAT	TCATCGGTCA	3217
Db	4975	GG	GTAAAC	CCCTTCT	CCAAAT	CAGACAA	ACGTTTCAA	AATCTTTTCTTCAT	TCATCGGTCA	4916
Qy	3218	TAA	ATCCGAT	CTCTTAC	AGGATATTT	TCAGTTT	CGTCAAT	TGCGGAT	TGTATATCCG	3277
Db	4915	TAA	ATCCGAT	CTCTTAC	AGGATATTT	TCAGTTT	CGTCAAT	TGCGGAT	TGTATATCCG	4856
Qy	3278	A	TTTATAT	TTTTTTT	TCGGTAT	TTTTTTT	TATTA	AAAA	CGTCTCA	3318
Db	4855	A	TTTATAT	TTTTTTT	TCGGTAT	TTTTTTT	TATTA	AAAA	CGTCTTACA	4815

RESULT 4
US-10-462-128-3/c
; Sequence 3, Application US/10462128
; Publication No. US20040067579A1
; GENERAL INFORMATION:
; TITLE OF INVENTION: PROCESS FOR OXIDATION OF STEROIDS AND
; ;
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIERMAN & MUSERLIAN
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/462,128
; FILING DATE: 16-Jun-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: US/08/054,185

/	FILING DATE:	26-APR-1993
/	APPLICATION NUMBER:	US/08/002,608
/	FILING DATE:	11-JAN-1993
/	APPLICATION NUMBER:	US/07/474,857
/	FILING DATE:	30-OCT-1990
/	APPLICATION NUMBER:	US/07/474,798
/	FILING DATE:	16-JULY-1990
/	APPLICATION NUMBER:	PCI/NL89/00072
/	FILING DATE:	25-SEPT-1989
/	APPLICATION NUMBER:	NL/88/200904.6
/	FILING DATE:	06-MAY-88
/	APPLICATION NUMBER:	NL/88/202080.3
/	FILING DATE:	03-SEP-88
/	ATTORNEY/AGENT INFORMATION:	
/	NAME:	CHARLES A. MUSERLIAN
/	REGISTRATION NUMBER:	19,683
/	REFERENCE/DOCKET NUMBER:	146.1169 CON-1
/	TELECOMMUNICATION INFORMATION:	
/	TELEPHONE:	(212) 661-8000
/	TELEFAX:	(212) 661-8002
/	INFORMATION FOR SEQ ID NO:	3:
/	SEQUENCE CHARACTERISTICS:	
/	LENGTH:	7336
/	TYPE:	NUCLEIC ACID
/	STRANDEDNESS:	DOUBLE
/	TOPOLOGY:	UNKNOWN
/	FEATURE:	
/	NAME/KEY:	
/	LOCATION:	
/	IDENTIFICATION METHOD:	
/	OTHER INFORMATION:	PLASMID pBHA-1
/	SEQUENCE DESCRIPTION:	SEQ ID NO: 3:
/	US-10-462-128-3	
	Query Match	29.0%; Score 1493; DB 13; Length 7336;
	Best Local Similarity	98.1%; Pred.No.3.le-282;
	Matches 1511; Conservative	0; Mismatches 30; Indels 0; Gaps 0;
Qy	1778	TTCTTATCTTTGTAATAAAGGTAACTATTCCGGGATAGACTGTAAACATTTCCTCACGCATA 1837
Dd	5367	TATCTTTTTTAATATCCCAGCTGGCAATCGCGGATAGACTGTAAACATTTCCTCACGCATA 6308
Qy	1838	AAATCCCTTTTCATTTTCTAATGAATCATATATACCTTATATA TTAATCAATTTCGCTCAT 1897
Dd	6307	AAATCCCTTTTCATTTTCTAATGTAAATCATATACCTTATATTAATATCAATTTCGCTCAT 6248
Qy	1898	AANTAAATCCCTTTTCTTTATTACGCAAAATGCGCGGATTTAAGACACACCCCTTTATTCGGTT 1957
Dd	6247	AAATTAATCCCTTTTCTTTATTACGCAAAATGCGCGGATTTAAGACACACCCCTTTATTCGGTT 6188
Qy	1958	AATGCGGCCATGACAGCCATGATAATTAATAATCTAGGAGAAGTTAATAATACGTTAAC 2017
Dd	6187	AATGCGGCCATGACAGCCATGATAATTAATAATCTAGGAGAAGTTAATAATACGTTAAC 6128
Qy	2018	AACATGNTAATCAATATTAGAGTGATCGTTCAAATGCTATGCGTTTGTGACACATCCA 2077
Dd	6127	AACATGNTAATCAATATTAGAGTGATCGTTCAAATGCTATGCGTTTGTGACACATCCA 6068
Qy	2078	CTATATATCCGTTGCTGTTCTGTCCACTCCTCGAATCCCATTCACGAAATTCCTTAGCGGATT 2137
Dd	6067	CTATATATCCGTTGCTGTTCTGTCCACTCCTCGAATCCCATTCACGAAATTCCTTAGCGGATT 6008
Qy	2138	CCGAAGTTTCTCAGATCGGAAGTTGACCAGACATTACGAACCTGSCACAGATGGTTCAT 2197
Dd	6007	CCGAAGTTTCTCAGATCGGAAGTTGACCAGACATTACGAACCTGSCACAGATGGTTCAT 5948
Qy	2198	AACCTGAAGGAAGATCTGATTGCTTTACTGCTTCAAGTTAAAGACGGAAGCGCTCGTTCGTAT 2257
Dd	5947	AACCTGAAGGAAGATCTGATTGCTTTACTGCTTCAAGTTAAGACGGAAGCGCTCGTTCGTAT 5888
Qy	2258	AACAGATGCGATGATGACAGACCAATCAACATGGCACTCTGCGCATTTGCTACCTGTACAGTCA 2317
Dd	5887	AACAGATGCGATGATGACAGACCAATCAACATGGCACTCTGCGCATTTGCTACCTGTACAGTCA 5828

Qy	2318	AGGATCGGTAGAAATGTTGTCGGTCTCTTGACACAGGAATATTAGCCCATTTGCCCTGCATATT	2377
Db	5827	AGGATCGGTAGAAATGTTGTCGGTCTCTTGACACAGGAATATTAGCCCATTTGCCCTGCATATT	5768
Qy	2378	CAAAACAGCTCTTCTACGATAAAGGCACAAATCGCATCGTGGAAACGTTTGGGCTTCTACCG	2437
Db	5767	CAAAACAGCTCTTCTACGATAAAGGCACAAATCGCATCGTGGAAACGTTTGGGCTTCTACCG	5708
Qy	2438	ATTTAGCAGTTGTATACATCTTCTCTAAGTATCCACCTGAAATCATATAATCGCCAAATAG	2497
Db	5707	ATTTAGCAGTTGTATACATCTTCTCTAAGTATCCACCTGAAATCATATAATCGCCAAATAG	5648
Qy	2498	AGAAAAATTTGACCATGTGTAAAGCGGCCAATCTGATTCACCTCGAGATGATATAATCTAGTA	2557
Db	5647	AGAAAAATTTGACCATGTGTAAAGCGGCCAATCTGATTCACCTCGAGATGATATAATCTAGTA	5588
Qy	2558	GAATCTCTTCGGTATCAAAATTCATCTCCACCTTCACCTCACCGGTTGTCCATTATCGC	2617
Db	5587	GAATCTCTTCGGTATCAAAATTCACCTTCACCTTCACCTCACCGGTTGTCCATTATCGC	5528
Qy	2618	TGAACCTCTGCTCTCTGTGACATGACACACATCATCTCAATATCCGAATAGGCGCCAT	2677
Db	5527	TGAACCTCTGCTCTCTGTGACATGACACACATCATCTCAATATCCGAATAGGCGCCAT	5468
Qy	2678	CAGCTGACGACCAAGAGAGCCATAAACAACCAATAGCCTTAACATCATGCCATATTTAT	2737
Db	5467	CAGCTGACGACCAAGAGAGCCATAAACAACCAATAGCCTTAACATCATGCCATATTTAT	5408
Qy	2738	CCAATATTCGTTCCCTTAAATTTATGAACCAATCTTCATCTTCTCTCTAGTCATTATTA	2797
Db	5407	CCAATATTCGTTCCCTTAAATTTATGAACCAATCTTCATCTTCTCTCTAGTCATTATTA	5348
Qy	2798	TGCGTCCATCACTATTCCTATTCCTCTTTTCAGATAAATTTAGATTTGCTTTTCTAAATA	2857
Db	5347	TGCGTCCATCACTATTCCTCTCTTTTCAGATAAATTTAGATTTGCTTTTCTAAATA	5288
Qy	2858	AGAAATATTTGGAGACGACCGTCTCTATTCAGCTATTATTAACCTCGTCTCTCTAAGCATCC	2917
Db	5287	AGAAATATTTGGAGACGACCGTCTCTATTCAGCTATTATTAACCTCGTCTCTCTAAGCATCC	5228
Qy	2918	TTCAATCCCTTTTAATAACAATATTAGCATCTAATCTTCAACAAACTGGCCCGGTTGTGA	2977
Db	5227	TTCAATCCCTTTTAATAACAATATTAGCATCTAATCTTCAACAAACTGGCCCGGTTGTGA	5168
Qy	2978	ACTACTCTTTAATAAATAATTTTTCGGTTCCTCAATTCACATTTGCAATTAATAGAAAAATC	3037
Db	5167	ACTACTCTTTAATAAATAATTTTTCGGTTCCTCAATTCACATTTGCAATTAATAGAAAAATC	5108
Qy	3038	CATCTTCATCGGCTTTTCGTGCATCATCTGTATGTAATGAAATCGCCCTTCTCTGTGTCAAT	3097
Db	5107	CATCTTCATCGGCTTTTCGTGCATCATCTGTATGTAATGAAATCGCCCTTCTCTGTGTCAAT	5048
Qy	3098	CAAGGTTTAAATTTTATGTATTCTTTTAAACAAACCAATAGGAGATTAAACCTTTTAC	3157
Db	5047	CAAGGTTTAAATTTTATGTATTCTTTTAAACAAACCAATAGGAGATTAAACCTTTTAC	4988
Qy	3158	GGTGTAACCTTCCTCCAAATCAGACAAACGGTTTCAAAATCTTTTCTTCATCATCGGTCA	3217
Db	4987	GGTGTAACCTTCCTCCAAATCAGACAAACGGTTTCAAAATCTTTTCTTCATCATCGGTCA	4928
Qy	3218	TAAAAATCCGTATCCTTTACAGGATATTTTTCGAGTTTTCGTCAAATGCGGATGTATATCCG	3277
Db	4927	TAAAAATCCGTATCCTTTACAGGATATTTTTCGAGTTTTCGTCAAATGCGGATGTATATCCG	4868
Qy	3278	ATTATATATTAATTTTCGGTATTTTTTTATATAAAACGCTCTCA	3318
Db	4867	ATTATATATTAATTTTCGGTATTTTTTTATATAAAACGCTCTCA	4827

RESULT 5
US-10-084-814-3/c
; Sequence 3, Application US/10084814

Publication No.	US20030108982A1	6187
GENERAL INFORMATION:		
APPLICANT:	SLIJKHUIS, HERMAN; SELTEN, GERARDUS CORNELIS MARIA; SMAAL, ERIC BASTIAAN	
TITLE OF INVENTION:	PROCESS FOR OXIDATION OF STERIODS AND GENETICALLY ENGINEERED CELLS USED THEREIN	
NUMBER OF SEQUENCES:	79	
CORRESPONDENCE ADDRESS:		
ADDRESS:	BIERMAN, MUSERLIAN & LUCAS	
STREET:	600 THIRD AVENUE	
CITY:	NEW YORK	
STATE:	NEW YORK	
COUNTRY:	USA	
ZIP:	10016	
COMPUTER READABLE FORM:		
MEDIUM TYPE:	FLOPPY DISK	
COMPUTER:	IBM PC COMPATIBLE	
OPERATING SYSTEM:	PC-DOS/MS-DOS	
SOFTWARE:	MICROSOFT WORD 97	
CURRENT APPLICATION DATA:		
APPLICATION NUMBER:	US/10/084,814	
FILING DATE:	26-Feb-2002	
PRIOR APPLICATION DATA:		
APPLICATION NUMBER:	08/418,085	
FILING DATE:	06-APR-1995	
APPLICATION NUMBER:	08/054,185	
FILING DATE:	26-APR-1993	
APPLICATION NUMBER:	08/002,608	
FILING DATE:	11-JAN-1993	
APPLICATION NUMBER:	07/474,857	
FILING DATE:	30-OCT-1990	
APPLICATION NUMBER:	07/474,798	
FILING DATE:	16-JULY-1990	
APPLICATION NUMBER:	PCT/NL89/00072	
FILING DATE:	25-SEPT-1989	
APPLICATION NUMBER:	NL88/200904.6	
FILING DATE:	06-MAY-1988	
APPLICATION NUMBER:	NL/88/202080.3	
FILING DATE:	03-SEP-1988	
ATTORNEY/AGENT INFORMATION:		
NAME:	CHARLES A. MUSERLIAN	
REGISTRATION NUMBER:	19,683	
REFERENCE/DOCKET NUMBER:	146.1169- CON-1-DIV-1	
TELECOMMUNICATION INFORMATION:		
TELEPHONE:	(212) 661-8000	
TELEFAX:	(212) 661-8002	
INFORMATION FOR SEQ ID NO:	3:	
SEQUENCE CHARACTERISTICS:		
LENGTH:	7336 BASE PAIRS	
TYPE:	NUCLEIC ACID	
STRANDEDNESS:	DOUBLE	
TOPOLOGY:	UNKNOWN	
FEATURE:		
OTHER INFORMATION:	PLASMID pBHA-1	
SEQUENCE DESCRIPTION:	SEQ ID NO: 3:	
US-10-084-814-3		
Query Match	29.0%; Score 1493; DB 15; Length 7336;	
Best Local Similarity	98.1%; Pred. No. 3.1e-282;	
Matches 1511; Conservative	0; Mismatches 30; Indels 0; Gaps 0;	
1778	TTCTTATCTGTGATAATGAGGTAACTATTGCGGGATAGACTGTAAACATTCTCAGCATA	1837
6367	TACTCTTTTAAATATCCCGACTGGCAATGCGGGATAGACTGTAAACATTCTCAGCATA	6308
1838	AAATCCCTTCATTTTCTAATGTAATCTATTACCTTATTATTAAATCAATTCGCTCAT	1897
6307	AAATCCCTTCATTTTCTAATGTAATCTATTACCTTATTATTAAATCAATTCGCTCAT	6248
1898	AATTAATCTTTTCTTATTAGCGAAATGGCCGATTTAAGCAACACCTTTATTCGGTT	1957

QY 3218 TAAATCCGATCGTTACAGGATATTTGAGTTTCGTCATTCGCGATTGTATATCCG 3277
DB 4927 TAAATCCGATCGTTACAGGATATTTGAGTTTCGTCATTCGCGATTGTATATCCG 4868
QY 3278 ATTATATATTTATTTTCGTTATTTTATTTAAACGCTCA 3318
DB 4867 ATTATATATTTATTTTCGTTGATTCATTTGAACTTTTACA 4827

RESULT 7
US-10-402-466A-7/c
; Sequence 7, Application US/10402466A
; Publication No. US20040028695A1
; GENERAL INFORMATION:
; APPLICANT: Park, Sukjoon
; APPLICANT: Giri, Lallan
; TITLE OF INVENTION: RECOMBINANT IMMUNOGENIC COMPOSITIONS AND METHODS FOR PROTECTING
; AGAINST LETHAL INFECTIONS FROM BACILLUS ANTHRACIS
; FILE REFERENCE: 18933-00005
; CURRENT APPLICATION NUMBER: US/10/402,466A
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: 60/372,152
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 34
; SEQ ID NO 7
; LENGTH: 8198
; TYPE: DNA
; ORGANISM: Bacillus anthracis
; FEATURE:
; NAME/KEY:
; LOCATION:
; OTHER INFORMATION: Expression vector for B. anthracis Protective Antigen (PA). Ent
; OTHER INFORMATION: sequence is shown since the vector sequence is different from the
; OTHER INFORMATION: of the pBP vectors. The PA coding sequence is from 3735 to 5942.

US-10-402-466A-7

Query Match 29.0%; Score 1493; DB 13; Length 8198;
Best Local Similarity 98.1%; Pred. No. 3.3e-282;
Matches 1511; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1778 TTCTTATCTTGATATAAGGTAAGTAACTATTGCGGGATAGACTGTAACTATCTCAGCGATA 1837
DB 3395 TACTCTTTTAAATATCCCGACTGGCAATGCCGATGAGCTGTAACATTTCTCAGCGATA 3336
QY 1838 AAATCCCTTTTCATTTCTTAATGTAATCTATTAATCTATTAATTAATCAATTCGCTCAT 1897
DB 3335 AAATCCCTTTTCATTTCTTAATGTAATCTATTAATCTATTAATTAATCAATTCGCTCAT 3276
QY 1898 AATTAATCTTTTCTTATTAAGCAAAATGCCGATTTAAGCACACCTTTATTCGGTT 1957
DB 3275 AATTAATCTTTTCTTATTAAGCAAAATGCCGATTTAAGCACACCTTTATTCGGTT 3216
QY 1958 AATGCCCATGACAGCCATGATAATTAATACTAATACTAGGAGAAAGTTAATAAATACGTAACC 2017
DB 3215 AATGCCCATGACAGCCATGATAATTAATACTAATACTAGGAGAAAGTTAATAAATACGTAACC 3156
QY 2018 AACATGATTAACAAATTTAGAGGTATCGTTCAAAATGGTATGGTTTGCACATCCA 2077
DB 3155 AACATGATTAACAAATTTAGAGGTATCGTTCAAAATGGTATGGTTTGCACATCCA 3096
QY 2078 CTATATATCGTGTCTGTCTGCTCACTCCCTGAATCCATCCAGAAATCTCTAGCGATT 2137
DB 3095 CTATATATCGTGTCTGTCTGCTCACTCCCTGAATCCATCCAGAAATCTCTAGCGATT 3036
QY 2138 CCAGAAAGTTTCTCAGATCGGAAAGTTGACCGACATTAACGACTGGCAACAGATGTCAT 2197
DB 3035 CCAGAAAGTTTCTCAGATCGGAAAGTTGACCGACATTAACGACTGGCAACAGATGTCAT 2976
QY 2198 AACCTGAAGCAAGATCTGATGCTTAATCTGCTTCACTTAAGACCGAGCGCTCGTCTGAT 2257
DB 2975 AACCTGAAGCAAGATCTGATGCTTAACTGCTTCACTTAAGACCGAGCGCTCGTCTGAT 2916
QY 2258 AACAGATGCGATGATGACAGACCAATCAACATGACGACCTGCCATTCCTGTACAGTCA 2317

RESULT 8

DB 2915 AACAGATGCGATGATGACAGACCAATCAACATGGCACCTGCCATTTACCTACGTACAGTCA 2856
QY 2318 AGGATGGTAGAATCTGTCGGTCTTCGACACGAAATATTAGCCATTTGCTCGCATATT 2377
DB 2855 AGGATGGTAGAATGTTGTCGGTCTTCGACACGAAATATTAGCCATTTGCTCGCATATT 2796
QY 2378 CAAACAGCTCTTTCTACGATTAAGGGCAAAATCGCATCGTGGAAACGTTTGGGCTTCTACCG 2437
DB 2795 CAAACAGCTCTTTCTACGATTAAGGGCAAAATCGCATCGTGGAAACGTTTGGGCTTCTACCG 2736
QY 2438 ATTAGCAGTTTGTATACACTTTCTTAAGTATCCACCTGAATCATAAATCGGCAAAATAG 2497
DB 2735 ATTAGCAGTTTGTATACACTTTCTTAAGTATCCACCTGAATCATAAATCGGCAAAATAG 2676
QY 2498 AGAAAATTTGACCATGTGTAAGCGGCCAATCTGATTTCCACCTGAGATGCATATCTAGTA 2557
DB 2675 AGAAAATTTGACCAATGTGTAAGCGGCCAATCTGATTTCCACCTGAGATGCATATCTAGTA 2616
QY 2558 GAATCTCTTGGCTATCAAAATTCATTTCCACCTTCCACTCAGCGTTGTCCATTCATGCG 2617
DB 2615 GAATCTCTTGGCTATCAAAATTCATTTCCACCTTCCACTCAGCGTTGTCCATTCATGCG 2556
QY 2618 TGAATCTCTGCTCTGTTGATGACACACATCATCTCAATATCCGAATAGGGGCCAT 2677
DB 2555 TGAATCTCTGCTCTGTTGATGACACACATCATCTCAATATCCGAATAGGGGCCAT 2496
QY 2678 CAGTCTGACGACCAAGAGAGCCATAAACACCAATAGCCTTAACATCATCCCCATATTTAT 2737
DB 2495 CAGTCTGACGACCAAGAGAGCCATAAACACCAATAGCCTTAACATCATCCCCATATTTAT 2436
QY 2738 CCAATATTCGTTCTTAATTTTCATGACCAATCTTCAATCTTCTCTCTAGTCAATATTA 2797
DB 2435 CCAATATTCGTTCTTAATTTTCATGACCAATCTTCAATCTTCTCTCTAGTCAATATTA 2376
QY 2798 TTGGTCCATTCATTTCTCAATTTCCCTTTTCAGATAAATTTAGATTTGCTTTTCTAAATA 2857
DB 2375 TTGGTCCATTCATTTCTCAATTTCCCTTTTCAGATAAATTTAGATTTGCTTTTCTAAATA 2316
QY 2858 AGAATTTGGAGAGACACCGTTCTTATTCAGCTATTAATAACCTGCTTCTCTAAGCATCC 2917
DB 2315 AGAATTTGGAGAGACACCGTTCTTATTCAGCTATTAATAACCTGCTTCTCTAAGCATCC 2256
QY 2918 TTCAATCTTTTAAATAAATAAATTTTCGGTTCCCAATTTCCACATTAATAAGAAATC 2977
DB 2255 TTCAATCTTTTAAATAAATAAATTTTCGGTTCCCAATTTCCACATTAATAAGAAATC 2196
QY 2978 ACTACTCTTTTAAATAAATAAATTTTCGGTTCCCAATTTCCACATTAATAAGAAATC 3037
DB 2195 ACTACTCTTTTAAATAAATAAATTTTCGGTTCCCAATTTCCACATTAATAAGAAATC 2136
QY 3038 CATCTTTCATCGGCTTTTTCGTCATCATCTGTATGAATCAATCGCTTCTTCTGTGTCTAT 3097
DB 2135 CATCTTTCATCGGCTTTTTCGTCATCATCTGTATGAATCAATCGCTTCTTCTGTGTCTAT 2076
QY 3098 CAAGTTTAAATTTTATGTATTTCTTTTAAACAAACCAATAGGAGATTAACTTTTAC 3157
DB 2075 CAAGTTTAAATTTTATGTATTTCTTTTAAACAAACCAATAGGAGATTAACTTTTAC 2016
QY 3158 GGTGTAAACCTTCTCCAAATTCAGACAAAGCTTTCAAAATTTCTTTTCTTCTATCATCGGTCA 3217
DB 2015 GGTGTAAACCTTCTCCAAATTCAGACAAAGCTTTCAAAATTTCTTTTCTTCTATCATCGGTCA 1956
QY 3218 TAAATCCGATCTTTTACAGGATATTTTTCGAGTTTTCGTCATTCGTCATTTGTATATCCG 3277
DB 1955 TAAATCCGATCTTTTACAGGATATTTTTCGAGTTTTCGTCATTCGTCATTTGTATATCCG 1896
QY 3278 ATTTATATTTATTTTTCGTTATTTTATTTTATTTAAACGCTCTCA 3318
DB 1895 ATTTATATTTATTTTTCGTTGGAATCATTTGAACTTTTACA 1855


```

; FILE REFERENCE: 9286.6CT
; CURRENT APPLICATION NUMBER: US/10/385,415
; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: US 09/936,028
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: PCT/EP00/01899
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: DE 19910102.7
; PRIOR FILING DATE: 1999-03-08
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 5302
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: p6021-CAT Expression vector
US-10-385-415-2

Query Match      28.9%; Score 1487.6; DB 16; Length 5302;
Best Local Similarity 99.1%; Pred.No.3e-281;
Matches 1496; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY      1809  CGGGATAGAGCTGAACATCTTCACGCATAAAATCCCCCTTTCATTTCTTAATGTAATCTA 1868
DB      2716  CGGGATAGAGCTGAACATCTTCACGCATAAAATCCCCCTTTCATTTCTTAATGTAATCTA 2775

QY      1869  TTACCTTATTTAATTCAAATTCGCTCATAAATTAATCCCTTTCTTATTACGCAAAATCG 1928
DB      2776  TTACCTTATTTAATTCAAATTCGCTCATAAATTAATCCCTTTCTTATTACGCAAAATCG 2835

QY      1929  CCGGATTTAAGCACACCCCTTTATTTCGGTTAATCGGCATGACAGCCATGATAATTACTAA 1988
DB      2836  CCGGATTTAAGCACACCCCTTTATTTCGGTTAATCGGCATGACAGCCATGATAATTACTAA 2895

QY      1989  TACTAGAGAGATTAAATAATAGTAACCAACATGATTAAACAATTAATAGAGTGTCATCGT 2048
DB      2896  TACTAGAGAGATTAAATAATAGTAACCAACATGATTAAACAATTAATAGAGTGTCATCGT 2955

QY      2049  TCAAAATGGTATCGTTTTCGACACATCCACTATATATCCGGTGTCTGTCCACATCCCTG 2108
DB      2956  TCAAAAATGGTATCGTTTTCGACACATCCACTATATATCCGGTGTCTGTCCACATCCCTG 3015

QY      2109  AATCCCATTCCAGAAAATCTCTAGCGAATCCAGAAGTTTCTCAGAGTCGGAAGTTGACC 2168
DB      3016  AATCCCATTCCAGAAAATCTCTAGCGAATCCAGAAGTTTCTCAGAGTCGGAAGTTGACC 3075

QY      2169  AGACATTACGAATCGGCACAGATGGTCATAAACCCTGAAGAGAGATCTGATTCCTTAATCG 2228
DB      3076  AGACATTACGAATCGGCACAGATGGTCATAAACCCTGAAGAGAGATCTGATTCCTTAATCG 3135

QY      2229  TTCAGTTAAGACCGAAGCGCTCGTGTATAACACAGATCGATGATGACAGACCAATCAACAT 2288
DB      3136  TTCAGTTAAGACCGAAGCGCTCGTGTATAACACAGATCGATGATGACAGACCAATCAACAT 3195

QY      2289  GGCACCTGCCATTGCTACCTGTA CAGTCAAGGATGGTAGAAAATGTTCTCGGTGCTTTGCAC 2348
DB      3196  GGCACCTGCCATTGCTACCTGTA CAGTCAAGGATGGTAGAAAATGTTCTCGGTGCTTTGCAC 3255

QY      2349  ACGAATATTAGCCCAATTGGCTCGCATATTCAAA CAGCTCTTCTACGATAAAGGGCACAAT 2408
DB      3256  ACGAATATTAGCCCAATTGGCTCGCATATTCAAA CAGCTCTTCTACGATAAAGGGCACAAT 3315

QY      2409  CGCATCTGGAAACGTTTGGGCTTCTACGATTTAGCAGTTTGTATACACTTTCTCTAAGTA 2468
DB      3316  CGCATCTGGAAACGTTTGGGCTTCTACGATTTAGCAGTTTGTATACACTTTCTCTAAGTA 3375

QY      2469  TCCACCTGAATCATAAATTCGGCAAAATAGAGAAAAATTGACCATGTGTAGCGGCCAATC 2528
DB      3376  TCCACCTGAATCATAAATTCGGCAAAATAGAGAAAAATTGACCATGTGTAGCGGCCAATC 3435

QY      2529  TGATTTCCACCTCGAGATGCATAATCTAGTAGAATCTCTCGGTATCAAAATTCACCTCCAC 2588

```

RESULT 10
US-10-385-415-4
Sequence 4, Application US/10385415
Publication No. US20040014158A1
GENERAL INFORMATION:
APPLICANT: Bacter, Adelbert
APPLICANT: Fischer, Markus
TITLE OF INVENTION: PROTEIN CONJUGATES, METHOD, VECTORS, PROTEINS AND DNA FOR PRODUCING THEM, THEIR USE AND MEDICAMENTS AND VACCINES
TITLE OF INVENTION: A CERTAIN QUANTITY OF SAID PROTEIN CONJUGATES
FILE REFERENCE: 9286 6CT
CURRENT APPLICATION NUMBER: US/10/385.415
CURRENT FILING DATE: 2003-03-10
PRIOR APPLICATION NUMBER: US 09/936, 028
PRIOR FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: PCT/EP00/01899
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: DE 19910102.7

```
; PRIOR FILING DATE: 1999-03-08
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 5767
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: p602-BS-luSy Expression plasmid
US-10-385-415-4

Query Match      28.9%; Score 1487.6; DB 16; Length 5767;
Best Local Similarity 99.1%; Pred. No. 3.1e-281;
Matches 1496; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1809 CGGATAGACGTACATCTTCAGCAGTAAATCCCTTTTCATTTCTTAATGTAAATCTA 1868
DB 3191 CGGATAGACGTACATCTTCAGCAGTAAATCCCTTTTCATTTCTTAATGTAAATCTA 3240
QY 1869 TTACCTTATTATTAATCAATTCCTCATATTAATCTTTTCTTATTAACGAAATGG 1928
DB 3241 TTACCTTATTATTAATCAATTCCTCATATTAATCTTTTCTTATTAACGAAATGG 3300
QY 1929 CCCGATTTAAGCAGACCCCTTTATTCGGTTAATGGCCATGACGCCATGATTAATCTAA 1988
DB 3301 CCCGATTTAAGCAGACCCCTTTATTCGGTTAATGGCCATGACGCCATGATTAATCTAA 3360
QY 1989 TACTAGGAGAGATTAATAAATACGTAAACCAATGATTAACAAATATTAGAGGTTCATCGT 2048
DB 3361 TACTAGGAGAGATTAATAAATACGTAAACCAATGATTAACAAATATTAGAGGTTCATCGT 3420
QY 2049 TCAAAATGGTATGGTTTGAACATCCATATATATATCCGTGCTGTCGTCACTCTTG 2108
DB 3421 TCAAAATGGTATGGTTTGAACATCCATATATATATCCGTGCTGTCGTCACTCTTG 3480
QY 2109 AATCCCATTCAGAAATTCCTAGCGATTCCAGAAATTCCTAGAGTCGGAAGTTGACC 2168
DB 3481 AATCCCATTCAGAAATTCCTAGCGATTCCAGAAATTCCTAGAGTCGGAAGTTGACC 3540
QY 2169 AGCAATTCAGAAATTCCTAGCGATTCCAGAAATTCCTAGAGTCGGAAGTTGACC 2228
DB 3541 AGCAATTCAGAAATTCCTAGCGATTCCAGAAATTCCTAGAGTCGGAAGTTGACC 3600
QY 2229 TTCAGTTAAGACCGAAGCGCTGCTGATATACAGATGCGATGATGAGCAACCAATCAACAT 2288
DB 3601 TTCAGTTAAGACCGAAGCGCTGCTGATATACAGATGCGATGATGAGCAACCAATCAACAT 3660
QY 2289 GGCACCTGCCATTCCTACCTGATACAGTCAAGGATGGTAGAAATGTTCTCGGTCTTGAC 2348
DB 3661 GGCACCTGCCATTCCTACCTGATACAGTCAAGGATGGTAGAAATGTTCTCGGTCTTGAC 3720
QY 2349 AGCAATTCAGAAATTCCTAGCGATTCCAGAAATTCCTAGAGTCGGAAGTTGACC 2408
DB 3721 AGCAATTCAGAAATTCCTAGCGATTCCAGAAATTCCTAGAGTCGGAAGTTGACC 3780
QY 2409 CGCATCGTGGAAACGTTTGGGCTTCTACCGATTAGCAGTTTGAATACACTTTCTCTAAGTA 2468
DB 3781 CGCATCGTGGAAACGTTTGGGCTTCTACCGATTAGCAGTTTGAATACACTTTCTCTAAGTA 3840
QY 2469 TCCACCTGAATCATTAATTCGGCAAAATAGAGAAAATGACCATGTGTAGCGGCCAATC 2528
DB 3841 TCCACCTGAATCATTAATTCGGCAAAATAGAGAAAATGACCATGTGTAGCGGCCAATC 3900
QY 2529 TGATTTCCACCTGAGATGATTAATCTAGTAGAATCTCTTCGCTATCAAAATTCACCTCCAC 2588
DB 3901 TGATTTCCACCTGAGATGATTAATCTAGTAGAATCTCTTCGCTATCAAAATTCACCTCCAC 3960
QY 2589 CTTCACCTACCGGTTGCCATTCATGGCTGAATCTGCTCTCTGTTGACATGACACA 2648
DB 3961 CTTCACCTACCGGTTGCCATTCATGGCTGAATCTGCTCTCTGTTGACATGACACA 4020
QY 2649 CATCATCTCAATATCCGAATAGGCGCCATCAGTCTGACGACCAAGAGAGCCATAAACACC 2708
DB 4021 CATCATCTCAATATCCGAATAGGCGCCATCAGTCTGACGACCAAGAGAGCCATAAACACC 4080
QY 2709 AATAGCCTTAACATCATCCCATATTTTATCAATATTCGTTTCTTAATTTTCATGAACAAT 2768
DB 4081 AATAGCCTTAACATCATCCCATATTTTATCAATATTCGTTTCTTAATTTTCATGAACAAT 4140
QY 2769 CTTCAATCTTTCTCTAGTCATTTATTTGGTCCATCTACTATTTCTCAITCCCTTTTC 2828
DB 4141 CTTCAATCTTTCTCTAGTCATTTATTTGGTCCATCTACTATTTCTCAITCCCTTTTC 4200
QY 2829 AGATTAATTTTAGATTTGCTTTTCTTAATAAGATAATTTGGAGAGCACCGTCTTATTACAG 2888
DB 4201 AGATTAATTTTAGATTTGCTTTTCTTAATAAGATAATTTGGAGAGCACCGTCTTATTACAG 4260
QY 2889 CTATTAATAAATCTCGTCTTCTTAAGCATCTTCAATCTTTTAAATAACAATTTATAGCATCT 2948
DB 4261 CTATTAATAAATCTCGTCTTCTTAAGCATCTTCAATCTTTTAAATAACAATTTATAGCATCT 4320
QY 2949 AATCTTCAACAACTGGCGGCTTTGTTGAATCTCTTTAATAAATAATTTTCCGTTTC 3008
DB 4321 AATCTTCAACAACTGGCGGCTTTGTTGAATCTCTTTAATAAATAATTTTCCGTTTC 4380
QY 3009 CCAATTCACATTCGCAATAAATAGAAAATCCATCTTTCATCGGCTTTTTCGTCATCATCTGT 3068
DB 4381 CCAATTCACATTCGCAATAAATAGAAAATCCATCTTTCATCGGCTTTTTCGTCATCATCTGT 4440
QY 3069 ATGAATCAAAATCGCTCTTCTGTCATCAAGGTTTAAATTTTATGATTTCTTTTAA 3128
DB 4441 ATGAATCAAAATCGCTCTTCTGTCATCAAGGTTTAAATTTTATGATTTCTTTTAA 4500
QY 3129 CAAACCCACATAGGAGATTAACCTTTTACGGTGTAAACCTTCTCCCAATCAGACAAACG 3188
DB 4501 CAAACCCACATAGGAGATTAACCTTTTACGGTGTAAACCTTCTCCCAATCAGACAAACG 4560
QY 3189 TTTCAAAATCTTTTCTCATCATCGGTCAATAAATCCGATCCGTTTACAGGATTTTTCG 3248
DB 4561 TTTCAAAATCTTTTCTCATCATCGGTCAATAAATCCGATCCGTTTACAGGATTTTTCG 4620
QY 3249 AGTTTCGTCATTCGCGATGCTATATCCGATTTATATTTTTCGTTATTTTATTA 3308
DB 4621 AGTTTCGTCATTCGCGATGCTATATCCGATTTATATTTTTCGTTATTTTATTA 4680
QY 3309 AAACGCTCA 3318
DB 4681 AACTTTTACA 4690

RESULT 11
US-09-869-855A-3/c
; Sequence 3, Application US/09869855A
; Publication No. US20030044940A1
; GENERAL INFORMATION:
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Method for Increasing Gene Copy Number
; FILE REFERENCE: 10028-204-US
; CURRENT APPLICATION NUMBER: US/09/869,855A
; CURRENT FILING DATE: 2001-07-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 5793
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-869-855A-3

Query Match      23.5%; Score 1207.6; DB 10; Length 5793;
Best Local Similarity 98.5%; Pred. No. 2.8e-226;
Matches 1219; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1813 ATAGACTGTAACTCTCAGCATAAATCCCTTTTCATTTTCTAATGAAATCTATTAC 1872
```


2069 AGAGATCTCGAGTTGGCGCGCAAAATCCCTTTTCATTTTCTAATGTAATCTATTAC 2010
1873 CTTATTATTAATCAATTCGCTCATATAATTAATCCCTTTTCTATTATACGCAAAATGGCCCG 1932
2009 CTTATTATTAATCAATTCGCTCATATAATTAATCCCTTTTCTATTATAGGAAATGGCCCG 1950
1933 ATTAAAGCACACCCCTTTATTCGGTTAAATGGCGCATGACAGGCATGATAATTAATACT 1992
1949 ATTAAAGCACACCCCTTTATTCGGTTAAATGGCGCATGACAGGCATGATAATTAATACT 1890
1993 AGGAGAAGTTAATAATACGTAACCAACATGATTAACCAATATTAGAGGTCATCGTTCAA 2052
1889 AGGAGAAGTTAATAATACGTAACCAACATGATTAACCAATATTAGAGGTCATCGTTCAA 1830
2053 AATGATGCGTTTGGACACATCCACTATATATCCGTGTCGTCTGTCCTCACTCCCTGAATC 2112
1829 AATGATGCGTTTGGACACATCCACTATATATCCGTGTCGTCTGTCCTCACTCCCTGAATC 1770
2113 CCAATCCAGAAATCTCTAGAGATTCAGAGATTTCTCAGAGTGGGAAGTTGACCCAGAC 2172
1769 CCAATCCAGAAATCTCTAGAGATTCAGAGATTTCTCAGAGTGGGAAGTTGACCCAGAC 1710
2173 ATTACGAATGCGACAGATGGTCATAACCTGAAGGAAGATCTGATTGCTTAACCTGCTTCA 2232
1709 ATTACGAATGCGACAGATGGTCATAACCTGAAGGAAGATCTGATTGCTTAACCTGCTTCA 1650
2233 GTTAAGACGGAAGCGCTGCTGTATATAACAGATGCGATGATGACAGCAATCAACATGGCA 2292
1649 GTTAAGACGGAAGCGCTGCTGTATATAACAGATGCGATGATGACAGCAATCAACATGGCA 1590
2293 CCGTGCAATGCTACCTGTACAGTCAAGGATGGTAGAAATGTTGTCGGTCCCTTGACACAGA 2352
1589 CCGTGCAATGCTACCTGTACAGTCAAGGATGGTAGAAATGTTGTCGGTCCCTTGACACAGA 1530
2353 ATATTACGCAATTCGCTGTCATATTAACACAGCTCTTCTACGATGAAGGCAAAATCGCA 2412
1529 ATATTACGCAATTCGCTGTCATATTAACACAGCTCTTCTACGATGAAGGCAAAATCGCA- 1470
2413 TCGTGGAAGCTTTGGCTTCTACCGATTTAGCAGTTTGTATACATCTTCTCTAAGTATCGCA 2472
1469 TCGTGGAAGCTTTGGCTTCTACCGATTTAGCAGTTTGTATACATCTTCTCTAAGTATCGCA 1410
2473 CCGTAATCATAAATCGGCAAAATAGAGAAATATGACCATGTGTGAAGCGGCAATCTGAT 2532
1409 CCGTAATCATAAATCGGCAAAATAGAGAAATATGACCATGTGTGAAGCGGCAATCTGAT 1350
2533 TCCACCTGAGATGCAATATCTAGTAGAATCTCTTCGCTATCAAAATTCACCTCCACCTTC 2592
1349 TCCACCTGAGATGCAATATCTAGTAGAATCTCTTCGCTATCAAAATTCACCTCCACCTTC 1290
2593 CACTCACCGGTTGTCCATTCATGGCTGAACCTCTGCTCTCTGTTGACATGACACACATC 2652
1289 CACTCACCGGTTGTCCATTCATGGCTGAACCTCTGCTCTCTGTTGACATGACACACATC 1230
2653 ATCTCAATATCGAATAGGCGCCATCAGTCTGACGACCAAGAGAGCCATTAACACCAATA 2712
1229 ATCTCAATATCGAATAGGCGCCATCAGTCTGACGACCAAGAGAGCCATTAACACCAATA 1170
2713 GCCTTAACATATCCCATATTTATTCCAATATTCGTTCTCTTAATTTATGACCAATCTTC 2772
1169 GCCTTAACATATCCCATATTTATTCCAATATTCGTTCTCTTAATTTATGACCAATCTTC 1110
2773 ATCTTTCTCTCTAGTCAATATTTATGTCATTCATTCATTCCTCCCTTTTCAGAT 2832
1109 ATCTTTCTCTCTAGTCAATATTTATGTCATTCATTCATTCCTCCCTTTTCAGAT 1050
2833 AATTTAGATTTGCTTTTCTAATTAAGATATTTGGAGGACCCGTTCTTATTCAGCTAT 2892
1049 AATTTAGATTTGCTTTTCTAATTAAGATATTTGGAGGACCCGTTCTTATTCAGCTAT 990
2893 TAATAACTCGTCTCTTAAGCATCTTCAATCCCTTTAATAACATTTATAGCATCTAATC 2952
989 TAATAACTCGTCTCTTAAGCATCTTCAATCCCTTTAATAACATTTATAGCATCTAATC 930

2953 TTCAACAACATGCGCCGTTTGTGAACTACTCTTTAATAAAATAATTTTCGTTCCCAA 3012
929 TTCAACAACATGCGCCGTTTGTGAACTACTCTTTAATAAAATAATTTTCGTTCCCAA 870
3013 TTCCACATTCGCAATATAGAAATCCATCTTCATCGGC 3050
869 TTCCACATTCGCAATATAGAAATCCATCTTCATCGTC 832

RESULT 12

US-09-869-855A-2/c
; Sequence 2, Application US/09869855A
; Publication No. US2003004940A1
; GENERAL INFORMATION:
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Method For Increasing Gene Copy Number
; FILE REFERENCE: 10028.204-US
; CURRENT APPLICATION NUMBER: US/09/869,855A
; CURRENT FILING DATE: 2001-07-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 5943
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE: Synthetic
; OTHER INFORMATION: Synthetic
US-09-869-855A-2

Query Match 23.5%; Score 1207.6; DB 10; Length 5943;
Best Local Similarity 98.5%; Pred. No. 2.8e-226;
Matches 1219; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

1813 ATGACGTGTAACATCTCAGCATAAATCCCTTTTCATTTCTAATGTAATCTATTAC 1872
2069 AGAGATCTCGAGTTGGCGCGCAAAATCCCTTTTCATTTCTAATGTAATCTATTAC 2010
1873 CTTATTATTAATCAATTCGCTCATATAATTAATCTTTTCTTATTACGCAAAATGGCCCG 1932
2009 CTTATTATTAATCAATTCGCTCATATAATTAATCTTTTCTTATTAGGAAATGGCCCG 1950
1933 ATTTAAGCACACCCCTTTATTCGGTTAAATGGCGCATGACAGGCATGATAATTAATACT 1992
1949 ATTTAAGCACACCCCTTTATTCGGTTAAATGGCGCATGACAGGCATGATAATTAATACT 1890
1993 AGGAGAAGTTAATAATACGTAACCAACATGATTAACCAATATTAGAGGTCATCGTTCAA 2052
1889 AGGAGAAGTTAATAATACGTAACCAACATGATTAACCAATATTAGAGGTCATCGTTCAA 1830
2053 AATGGTATGCGTTTGGACACATCCACTATATATCCGTTGCTGTTCTGTCCACTCCTGAATC 2112
1829 AATGGTATGCGTTTGGACACATCCACTATATATCCGTTGCTGTTCTGTCCACTCCTGAATC 1770
2113 CCATTCCAGAAATCTCTAGCGATTCCAGAAAGTTTCTCAGAGTCGGAAGTTGACCCAGAC 2172
1769 CCATTCCAGAAATCTCTAGCGATTCCAGAAAGTTTCTCAGAGTCGGAAGTTGACCCAGAC 1710
2173 ATTACGAATCGGACACAGATGGTCATTAACCTGAAGGAAGATCTGATTGCTTTAACTGCTTCA 2232
1709 ATTACGAATCGGACACAGATGGTCATTAACCTGAAGGAAGATCTGATTGCTTTAACTGCTTCA 1650
2233 GTTAAGACGGAAGCGCTGCTGTATTAACAGATGCGATGATGACAGCAATCAACATGGCA 2292
1649 GTTAAGACGGAAGCGCTGCTGTATTAACAGATGCGATGATGACAGCAATCAACATGGCA 1590
2293 CCGTGCAATGCTACCTGTACAGTCAAGGATGGTAGAAATGTTGTCGGTCCCTTGACACAGA 2352
1589 CCGTGCAATGCTACCTGTACAGTCAAGGATGGTAGAAATGTTGTCGGTCCCTTGACACAGA 1530
2353 ATATTACGCAATTCGCTGTCATATTAACACAGCTCTTCTAAGTATAGGCAAAATCGCA 2412
1529 ATATTACGCAATTCGCTGTCATATTAACACAGCTCTTCTAAGTATAGGCAAAATCGCA 1470

2413 TCCTGGACGTTTGGGCTTCTACCGATTTAGCAGTTTGATACACTTCTCTAAGTATCCA 2472
Db TCGTGGACGTTTGGGCTTCTACCGATTTAGCAGTTTGATACACTTCTCTAAGTATCCA 1410
2473 CCTGAATCATAAATCGGCAAAATAGAGAAAATGACCAATGTGTGAAGCGGCCAATCTGAT 2532
Db CCTGAATCATAAATCGGCAAAATAGAGAAAATGACCAATGTGTGAAGCGGCCAATCTGAT 1350
2533 TCCACCTGAGATGATATCTAGTAGAATCTCTTCGGTATCAAAATCACTTCCACCTTC 2592
Db TCCACCTGAGATGATATCTAGTAGAATCTCTTCGGTATCAAAATCACTTCCACCTTC 1290
2593 CACTCACCGGTTGTCATTTCATGGCTGAACCTCTGCTTCCCTCTGTGTGACATGACACATC 2652
Db CACTCACCGGTTGTCATTTCATGGCTGAACCTCTGCTTCCCTCTGTGTGACATGACACATC 1230
2653 ATCTCAATATCCGAATAGGGCCCATCAGTCTGACGACCAAGAGAGCCATAAACACCAATA 2712
Db ATCTCAATATCCGAATAGGGCCCATCAGTCTGACGACCAAGAGAGCCATAAACACCAATA 1170
2713 GCCTTAACATATCCCAATATTTATCCAAATATTCGTTCCCTTAATTCATGAACAATCTTC 2772
Db GCCTTAACATATCCCAATATTTATCCAAATATTCGTTCCCTTAATTCATGAACAATCTTC 1110
2773 ATCTTCTCTCTAGTCAATATTAATGGTCCATCTCATCTATCTCTCTCTCTCTCTCT 2832
Db ATCTTCTCTCTAGTCAATATTAATGGTCCATCTCATCTATCTCTCTCTCTCTCTCTCT 1050
2833 AATTTAGATTTGCTTTCTTAATAGNATATTTGGAGACACCGTCTTATTCAGCTAT 2892
Db AATTTAGATTTGCTTTCTTAATAGNATATTTGGAGACACCGTCTTATTCAGCTAT 990
2893 TAATAACTCGTCTTCTTAAGCATCTTCAATCCCTTTTAATAACAATATATAGCATCTAATC 2952
Db TAATAACTCGTCTTCTTAAGCATCTTCAATCCCTTTTAATAACAATATATAGCATCTAATC 930
2953 TTCAACAACTGGCCGTTTGTGTAAGTCTTCTTAATAAATATTTTCCGTTCCCAA 3012
Db TTCAACAACTGGCCGTTTGTGTAAGTCTTCTTAATAAATATTTTCCGTTCCCAA 870
3013 TTCCCATTTGCAATATAGAAAATCCATCTTCATCGGC 3050
Db TTCCCATTTGCAATATAGAAAATCCATCTTCATCGTC 832

RESULT 13

US-10-032-393-16
; Sequence 16, Application US/10032393
; Publication No. US20030027286A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Wall, Daniel
; APPLICANT: Gross, Molly
; TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
; FILE REFERENCE: ELITRA.010A
; CURRENT APPLICATION NUMBER: US/10/032,393
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/259,434
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 6852
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Vector pXyl-T5 p15a

US-10-032-393-16

Query Match 19.1%; Score 981.2; DB 15; Length 6852;
Best Local Similarity 99.2%; Pred. No. 8.1e-182;
Matches 986; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
831 TTAGAAAACCGACTGTAAAGATGACGTCGGCATTCTCATATATATAAAGCCAGTCAT 890
Db TTAGAAAACCGACTGTAAAGATGACGTCGGCATTCTCATATATATAAAGCCAGTCAT 4602
891 TAGGCCTATCTGACAAATCTCGAATAGAGTTTCATAAACAATCCCTGCATGATAAACCATCAC 950
Db TAGGCCTATCTGACAAATCTCGAATAGAGTTTCATAAACAATCCCTGCATGATAAACCATCAC 4662
951 AAACAGAAATGATGTACCTGTAAAGATAGCGGTAAATATATTAATGAATACCTTTATTAATGA 1010
Db AAACAGAAATGATGTACCTGTAAAGATAGCGGTAAATATATTAATGAATACCTTTATTAATGA 4722
1011 ATTTTCCCTGCTAATAATGGTAGAAGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1070
Db ATTTTCCCTGCTAATAATGGTAGAAGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4782
4723 ATTTTCCCTGCTAATAATGGTAGAAGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4782
1071 ACCCAGTAATGAAGTCCATGGAAATATAGAAAGAAAAGCAATTTTCAGGTATAGGTG 1130
Db ACCCAGTAATGAAGTCCATGGAAATATAGAAAGAAAAGCAATTTTCAGGTATAGGTG 4842
4783 ACCCAGTAATGAAGTCCATGGAAATATAGAAAGAAAAGCAATTTTCAGGTATAGGTG 4842
1131 TTTTGGGAAACAAATTTCCCGAACCATTATATTTCTCTACATCAGAAAGGTATATAATCAT 1190
Db TTTTGGGAAACAAATTTCCCGAACCATTATATTTCTCTACATCAGAAAGGTATATAATCAT 4902
4843 TTTTGGGAAACAAATTTCCCGAACCATTATATTTCTCTACATCAGAAAGGTATATAATCAT 4902
1191 AAAAATCTTTTGAAGTCATTCTTTTACAGGAGTCCAAATACCGAGAAATCTTTTATAGATAC 1250
Db AAAAATCTTTTGAAGTCATTCTTTTACAGGAGTCCAAATACCGAGAAATCTTTTATAGATAC 4962
4903 AAAAATCTTTTGAAGTCATTCTTTTACAGGAGTCCAAATACCGAGAAATCTTTTATAGATAC 4962
1251 CATCAAAATTTGTAAGTGTCTTAATCTTATCCCAATACCTTAACCTCCGTCGCTAT 1310
Db CATCAAAATTTGTAAGTGTCTTAATCTTATCCCAATACCTTAACCTCCGTCGCTAT 5022
1311 TGTAAACAGTCTTAAAGCTGTATTTGAGTTTATCACCCTTTGTCTACTAAGAAAATAAATG 1370
Db TGTAAACAGTCTTAAAGCTGTATTTGAGTTTATCACCCTTTGTCTACTAAGAAAATAAATG 5082
1371 CAGGGTAAATTTATATCTCTCTTTTATGTTTGGGTATAAACACATATATCAATTT 1430
Db CAGGGTAAATTTATATCTCTCTTTTATGTTTGGGTATAAACACATATATCAATTT 5142
1431 CTGTGTTATCTAAAGTCGTTTGTGTTTCAAAATATATGATTAATATCTCTTTCTCT 1490
Db CTGTGTTATCTAAAGTCGTTTGTGTTTCAAAATATATGATTAATATCTCTTTCTCT 5202
1491 TCCAAATTTGTCTAAATCAATTTTATTAAGTTTCATTGTATGATGCTCCCTCTCTCTCTCT 1550
Db TCCAAATTTGTCTAAATCAATTTTATTAAGTTTCATTGTATGATGCTCCCTCTCTCTCTCT 5262
1551 AAGTCAATTTAGGAGGCTTACT 1610
Db AAGTCAATTTAGGAGGCTTACT 5322
1611 AGTCAATATTTACTGTAAACATAAATATATATTTTAAATAATATCCCACTTTATCCAAATTT 1670
Db AGTCAATATTTACTGTAAACATAAATATATATTTTAAATAATATCCCACTTTATCCAAATTT 5382
1671 GTTGTGTAACTAAATGGGTGCTTTAGTGAAGAAATAGAGACCAATTAATAAATGTCGTC 1730
Db GTTGTGTAACTAAATGGGTGCTTTAGTGAAGAAATAGAGACCAATTAATAAATGTCGTC 5442
1731 TTTTGTGTTTTTTTAAAGGATTTGAGCGTAGCGGAAAAATCCCTTTTCTTTCTTATCTGAT 1790
Db TTTTGTGTTTTTTTAAAGGATTTGAGCGTAGCGGAAAAATCCCTTTTCTTTCTTATCTGAT 5502
1791 AATAAGGTAACATTTGCGCGGATAGACTGTAAAC 1824
Db AATAAGGTAACATTTGCGCGGATAGACTGTAAAC 5536

RESULT 14
US-10-032-393-53
; Sequence 53, Application US/10032393
; Publication No. US20030027286A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Wall, Daniel
; APPLICANT: Gross, Molly
; TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
; FILE REFERENCE: ELITRA.010A
; CURRENT APPLICATION NUMBER: US/10/032.393
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/259,434
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 6852
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pXyl-T5-DD p15a
US-10-032-393-53

Query Match 19.1%; Score 981.2; DB 15; Length 6852;
Best Local Similarity 99.2%; Pred. No. 8.1e-181;
Matches 986; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 831 TTAGAAAACCGACTGTAAGAGTACAGTCGGCATTCATCTATATATATAAAGCCAGTCAT 890
DB 4543 TTAGAAAACCGACTGTAAGAGTACAGTCGGCATTCATCTATATATATAAAGCCAGTCAT 4602
QY 891 TAGGCCTATCTGACAAATTCCTGAATAGAGTTCATTAACAATCTGTCATGAACCAATCAC 950
DB 4603 TAGGCCTATCTGACAAATTCCTGAATAGAGTTCATTAACAATCTGTCATGAACCAATCAC 4662
QY 951 AAACAGAGTATGATACCTGTAAGAGTACAGTCGGCATTCATCTATATATATAAAGCCAGTCAT 1010
DB 4663 AAACAGAGTATGATACCTGTAAGAGTACAGTCGGCATTCATCTATATATATAAAGCCAGTCAT 4722
QY 1011 ATTTTCCTGCTGTAATATGAGTACAGTAAATTAATTAATTAATTAATTAATTAATTAATTA 1070
DB 4723 ATTTTCCTGCTGTAATATGAGTACAGTAAATTAATTAATTAATTAATTAATTAATTAATTA 4782
QY 1071 ACCAGTAAATGAAGTCCGTAAGAGTACAGTCGGCATTCATCTATATATATAAAGCCAGTCAT 1130
DB 4783 ACCAGTAAATGAAGTCCGTAAGAGTACAGTCGGCATTCATCTATATATATAAAGCCAGTCAT 4842
QY 1131 TTGGGAAACCAATTTCCCGAACCAATTTATTTCTTACATCAGAAAGGTATAAATCAT 1190
DB 4843 TTGGGAAACCAATTTCCCGAACCAATTTATTTCTTACATCAGAAAGGTATAAATCAT 4902
QY 1191 AAACCTCTTTGAAGTCAATTTTACAGGAGTCCAAATACAGAGATGTTTATAGATACAC 1250
DB 4903 AAACCTCTTTGAAGTCAATTTTACAGGAGTCCAAATACAGAGATGTTTATAGATACAC 4962
QY 1251 CATCAAAATGATTAAGTGGCTTAATCTTATCCCAATTAATCTCCGTCGCTAT 1310
DB 4963 CATCAAAATGATTAAGTGGCTTAATCTTATCCCAATTAATCTCCGTCGCTAT 5022
QY 1311 TGTAACAGTCTTAAAGCTGTATTTGAGTTATACACCTTGTCATCAAGAAATAAATG 1370
DB 5023 TGTAACAGTCTTAAAGCTGTATTTGAGTTATACACCTTGTCATCAAGAAATAAATG 5082
QY 1371 CAGGTAATAATTTATATCTCTGTTTATGTTTCGGTATATAACACTAATATCAATTT 1430
DB 5083 CAGGTAATAATTTATATCTCTGTTTATGTTTCGGTATATAACACTAATATCAATTT 5142

RESULT 15
US-10-466-642-24/c
; Sequence 24, Application US/10466642
; Publication No. US20040048343A1
; GENERAL INFORMATION:
; APPLICANT: OmniGene BioProducts
; TITLE OF INVENTION: METHODS AND MICROORGANISMS FOR THE PRODUCTION OF 3-(2-
; TITLE OF INVENTION: HYDROXY-3-METHYL-BUTYRLAMINO)-PROPIONIC ACID (HMBPA)
; FILE REFERENCE: BGI-1469C
; CURRENT APPLICATION NUMBER: US/10/466,642
; CURRENT FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 3867
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: vector
US-10-466-642-24

Query Match 19.0%; Score 978.4; DB 13; Length 3867;
Best Local Similarity 99.9%; Pred. No. 2.1e-181;
Matches 979; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 831 TTAGAAAACCGACTGTAAGAGTACAGTCGGCATTCATCTATATATATAAAGCCAGTCAT 890
DB 2460 TTAGAAAACCGACTGTAAGAGTACAGTCGGCATTCATCTATATATATAAAGCCAGTCAT 2401
QY 891 TAGGCCTATCTGACAAATTCCTGAATAGAGTTCATTAACAATCTGTCATGAATACCATCAC 950
DB 2400 TAGGCCTATCTGACAAATTCCTGAATAGAGTTCATTAACAATCTGTCATGAATACCATCAC 2341
QY 951 AAACAGAGTATGATACCTGTAAGAGTACAGTCGGCATTCATTAATTAATTAATTAATTAATGA 1010
DB 2340 AAACAGAGTATGATACCTGTAAGAGTACAGTCGGCATTCATTAATTAATTAATTAATTAATGA 2281
QY 1011 ATTTTCCTGCTGTAATATGAGTACAGTAAATTAATTAATTAATTAATTAATTAATTAATTA 1070
DB 2280 ATTTTCCTGCTGTAATATGAGTACAGTAAATTAATTAATTAATTAATTAATTAATTAATTA 2221
QY 1071 ACCAGTAAATGAAGTCCATGGAATATAAGAAAGAAAGCAATTTTCAGGTATAGGTG 1130
DB 2220 ACCAGTAAATGAAGTCCATGGAATATAAGAAAGAAAGCAATTTTCAGGTATAGGTG 2161

QY 1131 TTTTGGGAAACAATTTCCCGAACCAATATATTTCTCTACATCAGAAAGGTATAAATCAT 1190
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2160 TTTTGGGAAACAATTTCCCGAACCAATATATTTCTCTACATCAGAAAGGTATAAATCAT 2101
QY 1191 AAAACCTCTTGAAGTCATCTTTACAGGAGTCCAAATACACAGAGAAATGTTTTGATACAC 1250
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2100 AAAACCTCTTGAAGTCATCTTTACAGGAGTCCAAATACACAGAGAAATGTTTTGATACAC 2041
QY 1251 CATCAAAAATTTGATAAAGTGGCTCTAACTTTATCCCAATACCTAATCTCCGTCGCTAT 1310
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2040 CATCAAAAATTTGATAAAGTGGCTCTAACTTTATCCCAATACCTAATCTCCGTCGCTAT 1981
QY 1311 TGTAACCAAGTTCTAAAGCTGATTTGAGTTTATCACCCCTTGTCACTAAGAAATTAATG 1370
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1980 TGTAACCAAGTTCTAAAGCTGATTTGAGTTTATCACCCCTTGTCACTAAGAAATTAATG 1921
QY 1371 CAGGTAATAATTTATATCTTTCTTTGTTTTATGTTTCGGTATAAAACACTAATATCAATTT 1430
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1920 CAGGTAATAATTTATATCTTTCTTTGTTTTATGTTTCGGTATAAAACACTAATATCAATTT 1861
QY 1431 CTGTGGTTATACATAAAGTCGTTTGTGGTCAATAATGATTAATAATCTCTTTCTCT 1490
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1860 CTGTGGTTATACATAAAGTCGTTTGTGGTCAATAATGATTAATAATCTCTTTCTCT 1801
QY 1491 TCCAATTTGCTAAATCAATTTTATTAAGTTTCATTTGATATGCCCTCTAAATTTTATCT 1550
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1800 TCCAATTTGCTAAATCAATTTTATTAAGTTTCATTTGATATGCCCTCTAAATTTTATCT 1741
QY 1551 AAAGTGAATTTAGGAGGCTTACTTGTCTGCTTTCTTCAATAGAAATCAATCTTTTAA 1610
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1740 AAAGTGAATTTAGGAGGCTTACTTGTCTGCTTTCTTCAATAGAAATCAATCTTTTAA 1681
QY 1611 AGTCAATATTACTGTAACATAATATATATTTAAATATATCCACATTTCCAAATTTTC 1670
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1680 AGTCAATATTACTGTAACATAATATATATTTAAATATATCCACATTTCCAAATTTTC 1621
QY 1671 GTTCTTGAACATAATGGTGTCTTTAGTTGAAGAATAAAGACCAATTAATAAATGTGGTC 1730
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1620 GTTCTTGAACATAATGGTGTCTTTAGTTGAAGAATAAAGACCAATTAATAAATGTGGTC 1561
QY 1731 TTTTGTGTTTTTTAAAGGATTTGAGCGTAGCGAAATCCCTTTCTTTCTTATCTTGAT 1790
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1560 TTTTGTGTTTTTTAAAGGATTTGAGCGTAGCGAAATCCCTTTCTTTCTTATCTTGAT 1501
QY 1791 AATAAGGTAACATAATGGCG 1810
Db |||||||||||||||||||
1500 AATAAGGTAACATAATGGCG 1481

Search completed: September 24, 2004, 19:24:45
Job time : 1470.26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 23, 2004, 17:04:19 ; Search time 7801.56 Seconds
(without alignments)
19682.126 Million cell updates/sec

Title: US-10-030-390-1
Perfect score: 5142
Sequence: 1 gaattcgagctcggtaccg.....ccgtctgtagcttccttaag 5142

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_nam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rtd:*
- 26: em_gss_pbg:*
- 27: em_gss_vrl:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	200.2	3.9	456	9 AA060105	AA060105 mj71f11.r
2	200.2	3.9	456	14 WB3072	WB3072 mf09c04.r1
3	198.6	3.9	354	13 BX102466	BX102466 BX102466
4	198.6	3.9	390	9 AI323443	AI323443 mj71f11.x

5	198.6	3.9	451	9 AI893383	AI893383 mk16b08.y
6	198.6	3.9	455	9 AI323089	AI323089 mj71f11.y
7	198.6	3.9	473	13 BY708773	BY708773 BY708773
8	198.6	3.9	474	11 AK008625	AK008625 Mus muscu
9	198.6	3.9	474	11 AK010179	AK010179 Mus muscu
10	198.6	3.9	475	11 AK008809	AK008809 Mus muscu
11	198.6	3.9	477	11 AK008873	AK008873 Mus muscu
12	198.6	3.9	486	13 BY708705	BY708705 BY708705
13	198.6	3.9	505	13 BY708809	BY708809 BY708809
14	198.6	3.9	994	9 AV072530	AV072530 AV072530
15	195.4	3.8	384	13 BY707652	BY707652 BY707652
16	195.4	3.8	354	13 BX101861	BY101861 BY101861
17	185.8	3.6	311	10 BB564851	BB564851 BB564851
18	182.6	3.6	435	9 AA097254	AA097254 mk16b08.r
19	181	3.5	311	10 BB564891	BB564891 BB564891
20	180.2	3.5	443	14 W12684	W12684 ma52b04.r1
21	165	3.2	238	29 AY419513	AY419513 Mus muscu
22	163	3.2	307	10 BB565090	BB565090 BB565090
23	161.2	3.1	306	10 BB565053	BB565053 BB565053
24	156.4	3.0	332	9 AV076259	AV076259 AV076259
C 25	155	3.0	705	10 BB667664	BB667664 BB667664
C 26	155	3.0	1515	11 AK050258	AK050258 Mus muscu
27	154.8	3.0	263	10 BB564847	BB564847 BB564847
28	148.4	2.9	300	10 BB564905	BB564905 BB564905
C 29	141	2.7	421	10 BF707544	BF707544 AL489706 T. brucei
30	137.8	2.7	354	10 BF707544	BF707544 AL489706 T. brucei
31	131.8	2.6	264	10 BB564882	BB564882 BB564882
32	130.8	2.5	298	10 BB565023	BB565023 BB565023
33	130.8	2.5	812	28 BM536445	BM536445 5'AD3 Vad
34	130.4	2.5	602	28 B06973	B06973 CSRL-8b6-t
35	127.4	2.5	242	10 BB565149	BB565149 BB565149
36	125.2	2.4	235	10 BB564787	BB564787 BB564787
37	124.8	2.4	317	9 AV075620	AV075620 AV075620
38	122.6	2.4	320	9 AV081837	AV081837 AV081837
39	120.4	2.3	254	10 BB565113	BB565113 BB565113
40	120.2	2.3	499	29 TA301803Q	TA301803Q T. brucei
41	120.2	2.3	550	29 TA144305Q	TA144305Q T. brucei
42	119.8	2.3	304	9 AV075572	AV075572 AV075572
43	118	2.3	503	29 TA344H03Q	TA344H03Q T. brucei
44	117.4	2.3	653	29 TA371H07Q	TA371H07Q T. brucei
45	117.4	2.3	659	29 TA371A11Q	TA371A11Q T. brucei

ALIGNMENTS

RESULT 1
AA060105
LOCUS
DEFINITION
AA060105 456 bp mRNA linear EST 23-SEP-1996
mj71f11.r1 Soares mouse p3NNF19.5 Mus musculus cDNA clone
IMAGE:481581 5' similar to gb:221858 M.Musculus mRNA for p domain
protein (MOUSE);, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AA060105
AA060105.1 GI:1553793
EST.
Mus musculus (house mouse)

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 456)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

TITLE
JOURNAL
COMMENT

The WashU-HMMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:292325
 Putative full length read
 vector to vector length is 457
 Seq primer: -28M13 rev2 from Amersham.
 Location/Qualifiers
 1..456

FEATURES

source
 1..456
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:481581"
 /dev_stage="19.5 dpc total fetus"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares mouse p3NM19.5"
 /note="Vector: p7T3D (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5],
 TGTACCAATCTGAAGTGGAGCGCGCATTTTCTTTT 3',
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified p7T3 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M.Fatima Bonaldo. RNA was kindly provided by
 Dr. Minoru Ko (Wayne State University)."

ORIGIN

Query Match 3.9%; Score 200.2; DB 9; Length 456;
 Best Local Similarity 88.6%; Pred. No. 5.4e-26;
 Matches 217; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
 QY 153 TGTCTACAGTCATATTCTGCTGCAGCCCGTGTGTCAGGTGTTTACGCCAGGCCCCAGG 212
 Db |||||
 31 TGTGTCCTCGCTGTGGTCTCATGTGGCTTCGAGCCTTCCAGCCAGGCCCCAGG 90
 QY 213 CCCAGGCCAGGAAGAAACATGTATCGCCCGCCCGGAGAGGATAAATTGGCTTCC 272
 Db |||||
 91 CCCAGGCCAGGAAGAAACATGTATCGCCCGCCCGGAGAGGATAAATTGGCTTCC 150
 QY 273 CCGGTGTACCGCCAGCAGTGCACGAGAGAGGTGCTGTTTGTATGACAGTGTCCGG 332
 Db |||||
 151 CCGGTGTACCGCCAGCAGTGCACGAGAGAGGTGCTGTTTGTATGACAGTGTCCGG 210
 QY 333 GATTCCTCGTGTGCTTCCACCCCATGGCCATCGAGAACACTCAAGAGAGAAATGTCCT 392
 Db |||||
 211 GATTCCTCGTGTGCTTCCACCCCATGGCCATCGAGAACACTCAAGAGAGAAATGTCCT 270
 QY 393 TCTAA 397
 Db |||||
 271 TCTAA 275

RESULT 2

W83072 456 bp mRNA linear EST 12-SEP-1996
 LOCUS m09c04.r1 Soares mouse p3NM19.5 Mus musculus cDNA clone
 DEFINITION IMAGE:404550 5', similar to gb:221858 M.Musculus mRNA for P domain
 protein (MOUSE);, mRNA sequence.

ACCESSION

W83072

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 456)
 AUTHOR Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.

TITLE

JOURNAL

COMMENT

The WashU-HHMI Mouse EST Project
 Unpublished (1996)
 On Sep 12, 1996 this sequence version replaced gi:1394024.
 Contact: Marra M/Mouse EST Project
 WashU-HHMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:248318
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 321.
 Location/Qualifiers

FEATURES

source

1..456
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:404550"
 /dev_stage="19.5 dpc total fetus"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares mouse p3NM19.5"
 /note="Vector: p7T3D (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5],
 TGTACCAATCTGAAGTGGAGCGCGCATTTTCTTTT 3',
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified p7T3 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M.Fatima Bonaldo. RNA was kindly provided by
 Dr. Minoru Ko (Wayne State University)."

ORIGIN

Query Match 3.9%; Score 200.2; DB 14; Length 456;
 Best Local Similarity 88.6%; Pred. No. 5.4e-26;
 Matches 217; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
 QY 153 TGTCTACAGTCATATTCTGCTGCAGCCCGTGTGTCAGGTGTTTACGCCAGGCCCCAGG 212
 Db |||||
 31 TGTGTCCTCGCTGTGGTCTCATGTGGCTTCGAGCCTTCCAGCCAGGCCCCAGG 90
 QY 213 CCCAGGCCAGGAAGAAACATGTATCGCCCGCCCGGAGAGGATAAATTGGCTTCC 272
 Db |||||
 91 CCCAGGCCAGGAAGAAACATGTATCGCCCGCCCGGAGAGGATAAATTGGCTTCC 150
 QY 273 CCGGTGTACCGCCAGCAGTGCACGAGAGAGGTGCTGTTTGTATGACAGTGTCCGG 332
 Db |||||
 151 CCGGTGTACCGCCAGCAGTGCACGAGAGAGGTGCTGTTTGTATGACAGTGTCCGG 210
 QY 333 GATTCCTCGTGTGCTTCCACCCCATGGCCATCGAGAACACTCAAGAGAGAAATGTCCT 392
 Db |||||
 211 GATTCCTCGTGTGCTTCCACCCCATGGCCATCGAGAACACTCAAGAGAGAAATGTCCT 270
 QY 393 TCTAA 397
 Db |||||
 271 TCTAA 275

RESULT 3

LOCUS

DEFINITION

BY102466 354 bp mRNA linear EST 07-DEC-2002
 BY102466 RIKEN full-length enriched, pooled tissues, adult spleen,
 etc. Mus musculus cDNA clone K630145D13 5', mRNA sequence.

ACCESSION

BY102466

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

This clone was previously sequenced on the 5' end only, this new data is from the 3' end
High quality sequence stop: 376.

FEATURES

Location/Qualifiers
1..390
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:481581"
/dev_stage="19.5 dpc total fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares mouse p3NMF19.5"
/note="Vector: p7T3D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCCGCAATTTTCTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Patima Bonaldo. RNA was kindly provided by Dr. Minoru KO (Wayne State University)."

ORIGIN

Query Match 3.9%; Score 198.6; DB 9; Length 390;
Best Local Similarity 98.0%; Pred. No. 1.1e-25;
Matches 201; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 193 TGTTTACGCCAGGCCAGGCCAGGCCAGGAAACATGTATCATGCCCCCGGGA 252
DB 384 TGCCAGGCCAGGCCAGGCCAGGCCAGGAAACATGTATCATGCCCCCGGGA 325
QY 253 GAGGATAAATTGTGCTTCCCGGTGTCCAGGCCAGGAGTGCAGGAGAGTTGCTG 312
DB 324 GAGGATAAATTGTGCTTCCCGGTGTCCAGGCCAGGAGTGCAGGAGAGTTGCTG 265
QY 313 TTTTATGACAGTGTCCCGGTGTCCAGGCCAGGAGTGCAGGAGAGTTGCTG 372
DB 264 TTTTATGACAGTGTCCCGGTGTCCAGGCCAGGAGTGCAGGAGAGTTGCTG 205
QY 373 TCAGAAGAAGAATGTCCTTCTAA 397
DB 204 TCAGAAGAAGAATGTCCTTCTAA 180

RESULT 5

AI893383
LOCUS
DEFINITION
mk16b08.y1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
IMAGE:493047 5', similar to gb:221858 M.Musculus mRNA for P domain protein (MOUSE);, mRNA sequence.

ACCESSION

AI893383
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)

REFERENCE

AUTHORS
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 451)
Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)

TITLE

JOURNAL
COMMENT
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
This read is a RESEQUENCE of a previously sequenced mouse clone
This read has been verified (found to hit its original self in the correct orientation)

MGI:296495

Seq primer: -40RP from Gibco
High quality sequence stop: 448.

FEATURES

Location/Qualifiers
1..451
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:493047"
/dev_stage="19.5 dpc total fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares mouse p3NMF19.5"
/note="Vector: p7T3D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCCGCAATTTTCTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Patima Bonaldo. RNA was kindly provided by Dr. Minoru KO (Wayne State University)."

ORIGIN

Query Match 3.9%; Score 198.6; DB 9; Length 451;
Best Local Similarity 98.0%; Pred. No. 1.1e-25;
Matches 201; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 193 TGTTTACGCCAGGCCAGGCCAGGCCAGGAAACATGTATCATGCCCCCGGGA 252
DB 66 TGCCAGGCCAGGCCAGGCCAGGCCAGGAAACATGTATCATGCCCCCGGGA 125
QY 253 GAGGATAAATTGTGCTTCCCGGTGTCCAGGCCAGGAGTGCAGGAGAGTTGCTG 312
DB 126 GAGGATAAATTGTGCTTCCCGGTGTCCAGGCCAGGAGTGCAGGAGAGTTGCTG 185
QY 313 TTTTATGACAGTGTCCCGGTGTCCAGGCCAGGAGTGCAGGAGAGTTGCTG 372
DB 186 TTTTATGACAGTGTCCCGGTGTCCAGGCCAGGAGTGCAGGAGAGTTGCTG 245
QY 373 TCAGAAGAAGAATGTCCTTCTAA 397
DB 246 TCAGAAGAAGAATGTCCTTCTAA 270

RESULT 6

AI323089
LOCUS
DEFINITION
mj7f11.y1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
IMAGE:481581 5', similar to gb:221858 M.Musculus mRNA for P domain protein (MOUSE);, mRNA sequence.

ACCESSION

AI323089
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)

REFERENCE

AUTHORS
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 451)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lemmon,G., Soares,B., Wilson,R. and Waterston,R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)

TITLE

JOURNAL
COMMENT
Contact: Marra M/Mouse EST Project

WASHU-HHMI Mouse EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:292325
This read is a RESEQUENCE of a previously sequenced mouse clone
correct orientation)
Putative full length read
vector to vector length is 457
Seq primer: -40RP from Gibco
High quality sequence stop: 439.
Location/Qualifiers
1. 455
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:481581"
/dev_stages="DHI0B (ampicillin resistant)"
/lab_host="DHI0B (ampicillin resistant)"
/clone_lib="Seares mouse p3MMP19.5"
/note="Vector: p773D (Pharmacia) with a modified
polylinker; Site1: Not I; Site2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCCGCGCATTTTTTTTTTTTTT 3']
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p773 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Scares and M.Fatima Bonaldo. RNA was kindly provided by
Dr. Minoru Ko (Wayne State University)."

ORIGIN

Query Match 3.9%; Score 198.6; DB 9; Length 455;
Best Local Similarity 98.0%; Pred. No. 1.1e-25;
Matches 201; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 193 TGTTCACCCAGGCCAGGCCAGGCCAGGAGAAACATGTATCATGTCGCCCGGGA 252
DB 72 TGCCAGCCAGCCAGGCCAGGCCAGGCCAGGAGAAACATGTATCATGTCGCCCGGGA 131
QY 253 GAGGATTAATTTGCTTCCCGGTTCCACCCAGCAGTGCACGGAGAGGTTGCTG 312
DB 132 GAGGATAAATTTGCTTCCCGGTTCCACCCAGCAGTGCACGGAGAGGTTGCTG 191
QY 313 TTTTGTATGACAGTGTCCGGGATTCCTGTTGCTTCCACCCAGCAGTGCACGGAGAGGTTGCTG 372
DB 192 TTTTGTATGACAGTGTCCGGGATTCCTGTTGCTTCCACCCAGCAGTGCACGGAGAGGTTGCTG 251
QY 373 TCAAGAAGAAGATGTCCTCTTAA 397
DB 252 TCAAGAAGAAGATGTCCTCTTAA 276

RESULT 7

BY708773 473 bp mRNA linear EST 16-DEC-2002
LOCUS BY708773 RIKEN full-length enriched, adult male stomach Mus
DEFINITION musculus cDNA clone 2210409105 5', mRNA sequence.
ACCESSION BY708773
VERSION BY708773.1 GI:271119967
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 473)
AUTHORS Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,

Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Mogami, A.,
Schonbach, C., Gojohori, T., Baldarelli, R., Hill, D.P., Buit, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Carbolodi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Guscinich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawajiri, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
Ravasi, R., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Santelin, A., Schneider, C., Sempie, C.A., Setou, M., Shinada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Inotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
22354683
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenhiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp
URL: http://genome.gsc.riken.go.jp/
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,
Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,
Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,
Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,
Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,
Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,
Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipillar sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN,
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
FEATURES
1. 473
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="2210409105"
/sex="male"

```

/tissue type="stomach"
/dev stage="adult"
/lab host="SOLR"
/clone_lib="RIKEN full-length enriched, adult male
stomach"
/note="Site 1: XhoI; Site 2: SstI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'.
GAGAGAGAGCGCCGCACTCGAGTTTTTTTTTTTTTN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'.
GAGAGAGATTCGAGTTAATTAATTAATCCCCCCCCCC 3']. cDNA
was cleaved with XhoI and SstI."

ORIGIN
Query Match 3.9%; Score 198.6; DB 13; Length 473;
Best Local Similarity 98.0%; Pred. No. 1e-25;
Matches 201; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 193 TGTTCACCCGAGCCGAGCCGAGCCGAGGAAACATGTTATCATGTCGCCCGCCGGA 252
DB 89 TGCCGAGCCGAGCCGAGCCGAGCCGAGGAAACATGTTATCATGTCGCCCGCCGGA 148
QY 253 GAGGATAAATTGTGCTTCCCGGTGTCACCGCCGAGGTCGACGAGAGAGTTCTG 312
DB 149 GAGGATAAATTGTGCTTCCCGGTGTCACCGCCGAGGTCGACGAGAGAGTTCTG 208
QY 313 TTTGATGACAGTGTCGCGGATTCCTCGTGTGCTTCCACCCCATGCCATCGAGAAC 372
DB 209 TTTGATGACAGTGTCGCGGATTCCTCGTGTGCTTCCACCCCATGCCATCGAGAAC 268
QY 373 TCAGAAGAAGAGTCCCTCTAA 397
DB 269 TCAAGAAGAAGATGCTCCTCTAA 293

RESULT 8
AK008625 474 bp mRNA linear HTC 20-SEP-2003
LOCUS Mus musculus adult male stomach cDNA, RIKEN full-length enriched
DEFINITION library, clone:2200002M16 product:trefol factor 1, full insert
sequence.
ACCESSION AK008625
VERSION AK008625.1 GI:12842925
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1
Carninci, P. and Hayashizaki, Y.
TITL High-efficiency full-length cDNA cloning
JOURN Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
AUTHORS 2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalizaion and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
JOURN 20499374
MEDLINE 11042159
PUBMED
AUTHORS 3
Shibata, K., Itoh, M., Aizawa, K., Nagakawa, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

CDS

```

```

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Ogawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ogawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipipillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
PUBMED 11076861
AUTHORS 4
The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 474)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hata, A., Hayatsu, N., Hiramoto, K.,
Hirakawa, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
Shizuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan [E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216]
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5'. GAGAGAGAAGGATCCAGAGCTTTTTTTTTTTTTTN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. Second
strand cDNA was prepared with the primer adapter of sequence [5'.
GAGAGAGATTCGAGTTAATTAATTAATCCCCCCCCCC 3']. cDNA was cleaved
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SOLR.
FEATURES
Location/Qualifiers
1..474
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:2200002M16"
/db_xref="MGI:1898509"
/db_xref="taxon:10090"
/clone="2200002M16"
/sex="male"
/tissue type="stomach"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev stage="adult"
31..294
/note="unnamed protein product; putative
trefol factor 1 (MGI:88135)"
/codon_start=1
/protein_id="BAB25787.1"

```

```

/db xref="GI:112842926"
/translators="MEKVT,CVLAVLMLAFGSLAQQAQAAQEFTCMAPRERIN
GFGVTAQQCTERGCGCFDDSVRGFPFPHFPAIENTOEBCPF"

ORIGIN
Query Match          3.9%; Score 198.6; DB 11; Length 474;
Best Local Similarity 98.0%; Pred. No. le-25;
Matches 201; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      193 TGTTTACGCCAGCCAGCCAGCCAGCCAGCCAGGAAGAACAATGTATCATGGCCCCCGGGA 252
Db       90 TGCACAGCCAGCCAGCCAGCCAGCCAGCCAGGAAGAACAATGTATCATGGCCCCCGGGA 149

QY      253 GAGGATAAATTGGCTTCCTCCCGTGTCACGCCAGCAGTGCACGAGAGAGGTTGCTG 312
Db       150 GAGGATAAATTGGCTTCCTCCCGTGTCACGCCAGCAGTGCACGAGAGAGGTTGCTG 209

QY      313 TTTTGATGACAGTGTCCGGGATTCCTCGTGTGCTTCACCCCATGGCCATCGAGAACAC 372
Db       210 TTTTGATGACAGTGTCCGGGATTCCTCGTGTGCTTCACCCCATGGCCATCGAGAACAC 269

QY      373 TCAGAAGAAGAATGTCCTCTCTAA 397
Db       270 TCAGAAGAAGAATGTCCTCTCTAA 294


RESULT 9
AKO10179
LOCUS
DEFINITION
Mus musculus adult male tongue cDNA, RIKEN full-length enriched
library, clone:2310075107 product:trefoil factor 1, full insert
sequence.
ACCESSION
AKO10179
VERSION
AKO10179.1 GI:12845437
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 Carninci,P. and Hayashizaki,Y.
  Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  Normalization and subtraction of cap-trapper-selected cDNAs to
  prepare full-length cDNA libraries for rapid discovery of new genes
  Genome Res. 10 (10), 1617-1630 (2000)
2 Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  RIKEN integrated sequence analysis (RISA) system-384-format
  sequencing pipeline with 384 multicapillary sequencer
  Genome Res. 10 (11), 1757-1771 (2000)
3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
  Konno,H., Akiyama,J., Nishi,K., Kitsumai,T., Tashiro,H., Itoh,M.,
  Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
  Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kasaiwagi,K.,
  Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,S., Watanuki,M.,
  Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai.J.
  Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  RIKEN integrated sequence analysis (RISA) system-384-format
  sequencing pipeline with 384 multicapillary sequencer
  Genome Res. 10 (11), 1757-1771 (2000)
4 The RIKEN Genome Exploration Research Group Phase II Team and the
  FANTOM Consortium.
  Functional annotation of a full-length mouse cDNA collection
  Nature 409, 695-690 (2001)
5 The FANTOM Consortium and the RIKEN Genome Exploration Research
  AUTHORS

```

Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 474)

AUTHORS
 ADACHI,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
 Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
 Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
 Hirata,K., Hori,F., Ictani,K., Ishii,Y., Itoh,M., Izawa,M.,
 Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
 Koya,S., Kurahara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
 Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
 Saito,H., Saito,R., Sakai,K., Sakai,K., Sano,H., Sasaki,D.,
 Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
 Tejima,H., Tagami,M., Takahashi,A., Takahashi,F., Tanaka,T.,
 Tazaki,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K.,
 Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

TITLE
 Direct Submission

JOURNAL
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
 URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
 Fax:81-45-503-9216)

COMMENT
 Please visit our web site (http://genome.gsc.riken.go.jp/) for
 further details.
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues. first strand cDNA was primed with a primer
 [5' GAGAGAGAGAGATCGAAGCTCTTTTTTTTTTTTTTNN 3', cDNA was
 prepared by using thermostable thermo-activated reverse transcriptase
 and subsequently enriched for full-length by cap-trapper. Second
 strand cDNA was prepared with the primer adapter of sequence [5'
 GAGAGAGATCTCGACTTAATTAATATCCGCCCCCCC 3']. cDNA was cleaved
 with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
 Host: SOLR.

FEATURES
 source
 1..474
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="PANTOM DB:2310075107"
 /db_xref="MG1:1898578"
 /db_xref="taxon:10090"
 /clone="2310075107"
 /sex="male"
 /tissue_type="tongue"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 31..294
 /notes="unnamed protein product; putative
 trefoil factor 1 (MGD|MG1:88135)"
 /codon_start=1
 /protein_id="BAB26751.1"
 /db_xref="GI:12845438"
 /translation="MSHKVTCVLAVLMLAFGLAQAAQAAQAEETCINAPRERINC
 GPFVTAQCCTERGCCFDDSVRGFPWCFHPMAIENTQEBCPF"

CDS
 31..294
 /notes="unnamed protein product; putative
 trefoil factor 1 (MGD|MG1:88135)"
 /codon_start=1
 /protein_id="BAB26751.1"
 /db_xref="GI:12845438"
 /translation="MSHKVTCVLAVLMLAFGLAQAAQAAQAEETCINAPRERINC
 GPFVTAQCCTERGCCFDDSVRGFPWCFHPMAIENTQEBCPF"

ORIGIN
 Query Match 3.9%; Score 198.6; DB 11; Length 474;
 Best Local Similarity 98.0%; Pred. No. 1e-25;
 Matches 201; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 193 TGTTTACGCCAGGCCAGGCCAGGCCAGGAGAAACATGTATCTATGCCCCCGGGA 252
 |||
 90 TGCCAGGCCAGGCCAGGCCAGGCCAGGAGAAACATGTATCTATGCCCCCGGGA 149
 |||
 253 GAGATAAATTGGCTTCCCGGTGTCACGCCCCAGCAGTGCACGGAGAGAGGTTGCTG 312
 |||
 150 GAGATAAATTGGCTTCCCGGTGTCACGCCCCAGCAGTGCACGGAGAGAGGTTGCTG 209
 |||


```
1. .486
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="2210403C21"
/sex="male"
```



```
adult,tissue_type=amnion,sex=female),(dev_stage=10 days
neonate,tissue_type=brain,sex=mix),(dev_stage=10 days
neonate,tissue_type=thymus,sex=mix),(dev_stage=10 days
neonate,tissue_type=heart,sex=mix)"

ORIGIN
Query Match      3.8%; Score 195.4; DB 13; Length 384;
Best Local Similarity 97.1%; Pred. No. 4.2e-25;
Matches 199; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 193 TGTTTACGCCAGGCCAGGCCAGGCCAGGAAACATGTATCATGGCCCCCGGGA 252
    |||
Db 86 TGCCCCAGGCCAGGCCAGGCCAGGAAACATGTATCATGGCCCCCGGGA 145
QY 253 GAGGATAAATTGTGGCTTCCCGGTGTACCGCCAGCAGTGCACGGAGAGAGTTGCTG 312
    |||
Db 146 GAGGATAAATTGTGGCTTCCCGGTGTACCGCCAGCAGTGCACGGAGAGAGTTGCTG 205
QY 313 TTTTGATGACAGTGTCCGGGGATTCCCGTGGTGTTCACACCCCATGGCCATCGAGAACAC 372
    |||
Db 206 TTTTGATGACAGTGTCCGGGGATTCCCGTGGTGTTCACACCCCATGGCCATCGAGAACAC 265
QY 373 TCAAGAAGAAGAAATGTCCCTTCTAA 397
    |||
Db 266 TCAAGAAGAAGAAATGGTCTCTAA 290
```

Search completed: September 24, 2004, 09:58:24
Job time : 7806.56 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2004, 18:06:49 ; Search time 219.773 Seconds
(without alignments)
12984.105 Million cell updates/sec

Title: US-10-030-390-1

Perfect score: 5142

Sequence: 1 gaattcgagctcggtaccg.....ccgtctgtagcttctctaag 5142

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	2467.2	48.0	8119	1	US-08-460-343B-1 Sequence 1, Appli
C 2	2467.2	48.0	8119	1	US-08-398-028B-1 Sequence 1, Appli
C 3	2467.2	48.0	8119	2	US-08-504-265B-1 Sequence 1, Appli
C 4	1636.8	31.8	3792	2	US-08-992-334-1 Sequence 1, Appli
C 5	1636.8	31.8	3792	3	US-08-302-752-1 Sequence 2, Appli
C 6	1635.2	31.8	5234	2	US-08-992-334-2 Sequence 2, Appli
C 7	1635.2	31.8	5234	3	US-08-302-752-2 Sequence 2, Appli
C 8	1635.2	31.8	6722	2	US-08-992-334-3 Sequence 3, Appli
C 9	1635.2	31.8	6722	3	US-08-302-752-3 Sequence 3, Appli
C 10	1493	29.0	7336	3	US-08-418-085-3 Sequence 3, Appli
C 11	1493	29.0	7336	3	US-09-099-011A-3 Sequence 3, Appli
C 12	1493	29.0	7336	4	US-08-470-368-21 Sequence 21, Appli
C 13	1493	29.0	7336	4	US-09-098-877B-3 Sequence 3, Appli
C 14	1493	29.0	9144	3	US-08-556-978B-79 Patent No. 5304637
C 15	1491.4	29.0	7326	6	5304637-21 Sequence 14, Appli
C 16	1479	28.8	5609	4	US-09-313-677-14 Sequence 15, Appli
C 17	1479	28.8	5609	4	US-09-313-677-15 Sequence 15, Appli
C 18	1479	28.8	7026	4	US-09-313-677-20 Sequence 20, Appli
C 19	1479	28.8	7344	4	US-09-313-677-16 Sequence 16, Appli
C 20	1084.6	21.1	6169	4	US-08-875-154-2 Sequence 2, Appli
C 21	867	16.9	3156	4	US-09-633-927-1 Sequence 1, Appli
C 22	830.8	16.2	10216	2	US-08-875-154-1 Sequence 1, Appli
C 23	763.8	14.9	1200	2	US-08-642-045B-3 Sequence 3, Appli
C 24	763.8	14.9	1200	3	US-08-852-268-1 Sequence 1, Appli
C 25	667.8	13.0	1102	1	US-08-297-294A-2 Sequence 2, Appli
C 26	408	7.9	3792	2	US-08-992-334-1 Sequence 1, Appli
C 27	408	7.9	3792	3	US-08-302-752-1 Sequence 1, Appli

C 28	408	7.9	5234	2	US-08-992-334-2	Sequence 2, Appli
C 29	408	7.9	5234	3	US-08-302-752-2	Sequence 2, Appli
C 30	408	7.9	6722	2	US-08-992-334-3	Sequence 3, Appli
C 31	408	7.9	6722	3	US-08-302-752-3	Sequence 3, Appli
C 32	347.2	6.8	5390	5	PCT-US96-12545-5	Sequence 5, Appli
C 33	162	3.2	5502	4	US-09-702-705-785	Sequence 785, App
C 34	162	3.2	5502	4	US-09-736-457-785	Sequence 785, App
C 35	162	3.2	5502	4	US-09-614-124B-785	Sequence 785, App
C 36	162	3.2	5502	4	US-09-671-325-785	Sequence 785, App
C 37	162	3.2	5502	4	US-09-589-184-785	Sequence 785, App
C 38	151.8	3.0	5217	4	US-09-100-703A-25	Sequence 25, Appli
C 39	150.8	2.9	5230	4	US-09-100-703A-26	Sequence 26, Appli
C 40	148.4	2.9	2321	4	US-09-198-839-2	Sequence 2, Appli
C 41	148.4	2.9	2321	4	US-09-562-834-2	Sequence 2, Appli
C 42	148.4	2.9	2321	4	US-09-995-396-2	Sequence 3, Appli
C 43	148.4	2.9	2337	3	US-09-198-839-3	Sequence 3, Appli
C 44	148.4	2.9	2337	4	US-09-562-834-3	Sequence 3, Appli
C 45	148.4	2.9	2337	4	US-09-995-396-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-460-343B-1/c
; Sequence 1, Application US/08460343B
; Patent No. 5741664

GENERAL INFORMATION:
; APPLICANT: Marcus D. Ballinger and James A. Wells
; TITLE OF INVENTION: SUBSTITUTED VARIANTS CAPABLE OF CLEAVING
; TITLE OF INVENTION: SUBSTRATES CONTAINING DIBASIC RESIDUES
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA: US/08/460,343B
; APPLICATION NUMBER: US/08/460,343B
; FILING DATE: 01-Jun-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/398028
; FILING DATE: 03-mar-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rubinec, Jeffrey S.
; REGISTRATION NUMBER: 36,575
; REFERENCE/DOCKET NUMBER: P0936C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8228
; TELEFAX: 650/952-9981
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8119 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-460-343B-1

Query Match 48.0%; Score 2467.2; DB 1; Length 8119;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2472; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Cyt 831 TTGAAACCGCCTGTAAGTACCTCGGCATTATCTCATATATAAAGCCAGTCAT 890
DB 7847 TTGAAACCGCCTGTAAGTACCTCGGCATTATCTCATATATAAAGCCAGTCAT 7788

QY 891 TAGGCCTATCTGACAAATCCCTGAATAGAGTTTCATAAAACAATCCCTGCGATGATAACCATCAC 950
Db 7787 TAGGCCTATCTGACAAATCCCTGAATAGAGTTTCATAAAACAATCCCTGCGATGATAACCATCAC 7728
QY 951 AAAACAGATGATGATCTGCTGTAAGATAGCGGTAAATATATGATATACCTTTATTAATGA 1010
Db 7727 AAAACAGATGATGATCTGCTGTAAGATAGCGGTAAATATATGATATACCTTTATTAATGA 7668
QY 1011 ATTTTCCTGCTGTAATATGCTGTAAGATAGCGGTAAATATATGATATACCTTTATTAATGA 1070
Db 7667 ATTTTCCTGCTGTAATATGCTGTAAGATAGCGGTAAATATATGATATACCTTTATTAATGA 7608
QY 1071 ACCCAGTAATGATGATCTGCTGTAAGATAGCGGTAAATATATGATATACCTTTATTAATGA 1130
Db 7607 ACCCAGTAATGATGATCTGCTGTAAGATAGCGGTAAATATATGATATACCTTTATTAATGA 7548
QY 1131 TTTTCGGGAAACAAATTTCCCGAACCAATATATTTCTCTACATCAGAAAGGTATATAATCAT 1190
Db 7547 TTTTCGGGAAACAAATTTCCCGAACCAATATATTTCTCTACATCAGAAAGGTATATAATCAT 7488
QY 1191 AAAACCTTTGAGTCAATCTTTTACAGGAGTCCAAATACACAGAGATGTTTTAGATACAC 1250
Db 7487 AAAACCTTTGAGTCAATCTTTTACAGGAGTCCAAATACACAGAGATGTTTTAGATACAC 7428
QY 1251 CATCAAAAATTTGATATAAGTGGCTCTAACTTATCCCAATACCTAACTCTCCGTCGCTAT 1310
Db 7427 CATCAAAAATTTGATATAAGTGGCTCTAACTTATCCCAATACCTAACTCTCCGTCGCTAT 7368
QY 1311 TGTAAACCAAGTCTTAAAGCTGTATTTGAGTTTATCACCCCTGTCTCACTAAGAAAATAAATG 1370
Db 7367 TGTAAACCAAGTCTTAAAGCTGTATTTGAGTTTATCACCCCTGTCTCACTAAGAAAATAAATG 7308
QY 1371 CAGGTAATAATTTATATCTTTCTGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 1430
Db 7307 CAGGTAATAATTTATATCTTTCTGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 7248
QY 1431 CTGTGGTTATATACTAAAGTCTGTTTCTGTTTCAATAATGATTAATAATATCTCTTTCTCT 1490
Db 7247 CTGTGGTTATATACTAAAGTCTGTTTCTGTTTCAATAATGATTAATAATATCTCTTTCTCT 7188
QY 1491 TCCAAATGCTAAATCAATTTTATTAAGTTCATTTGATAGCTCCCTAAATTTTATCT 1550
Db 7187 TCCAAATGCTAAATCAATTTTATTAAGTTCATTTGATAGCTCCCTAAATTTTATCT 7128
QY 1551 AAAGTGAATTTAGGAGGCTTACTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1610
Db 7127 AAAGTGAATTTAGGAGGCTTACTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7068
QY 1611 AGTCAATATTAAGTGAATTAATAATATATTTTAAATATATCCCAATTTATCCCAATTTTC 1670
Db 7067 AGTCAATATTAAGTGAATTAATAATATATTTTAAATATATCCCAATTTATCCCAATTTTC 7008
QY 1671 GTTCTGTAATTAAGTGAATTAATAATATATTTTAAATATATCCCAATTTATCCCAATTTTC 1730
Db 7007 GTTCTGTAATTAAGTGAATTAATAATATATTTTAAATATATCCCAATTTATCCCAATTTTC 6948
QY 1731 TTTTGTGTTTTTTTAAAGGATTTGAGCGTAGCGAATAATCTCTTTTCTTTTCTTTATCTTGAT 1790
Db 6947 TTTTGTGTTTTTTTAAAGGATTTGAGCGTAGCGAATAATCTCTTTTCTTTTCTTTATCTTGAT 6888
QY 1791 AATAAGGTAATTTGCGGATAGACTGTAACATTTCTCGCATATAAATCCCTTTCA 1850
Db 6887 AATAAGGTAATTTGCGGATAGACTGTAACATTTCTCGCATATAAATCCCTTTCA 6828
QY 1851 TTTTCTAATGTAATATCTATTTACTTTATTTAATTTCAATTTGCTCTCAATAATTAATCCCTTTT 1910
Db 6827 TTTTCTAATGTAATATCTATTTACTTTATTTAATTTCAATTTGCTCTCAATAATTAATCCCTTTT 6768
QY 1911 TCTTATTAAGCAAAATGGCCGATTTAAGCAGCACCCCTTTATTCGTTAATGCGGCATGAC 1970
Db 6767 TCTTATTAAGCAAAATGGCCGATTTAAGCAGCACCCCTTTATTCGTTAATGCGGCATGAC 6708

QY 1971 AGCCATGATAATTAATACTAGGAGGTTAATAAATAGCTACCAACATGATTACAA 2030
Db 6707 AGCCATGATAATTAATACTAGGAGGTTAATAAATAGCTACCAACATGATTACAA 6648
QY 2031 ATTATTAGAGGTCAATCGTTTCAAAATGGTATGCTTTTGACACATCCACTATATATCCGTTG 2090
Db 6647 ATTATTAGAGGTCAATCGTTTCAAAATGGTATGCTTTTGACACATCCACTATATATCCGTTG 6588
QY 2091 TGTCTCTGTCCTCTGTAATCCCAATTCAGAAATTTCTAGCGATTTCCAGAGTTTCTC 2150
Db 6587 TGTCTCTGTCCTCTGTAATCCCAATTCAGAAATTTCTAGCGATTTCCAGAGTTTCTC 6528
QY 2151 AGAGTCGGAAGTTGACACAGATTTACCAACTGGCAGAGATGGTCAATACCTCAAGGAAG 2210
Db 6527 AGAGTCGGAAGTTGACACAGATTTACCAACTGGCAGAGATGGTCAATACCTCAAGGAAG 6468
QY 2211 ATCTGATTGCTTAACTGCTTCAAGTTAAGACCGAGGCTGCTGATTAACAGATGCGATG 2270
Db 6467 ATCTGATTGCTTAACTGCTTCAAGTTAAGACCGAGGCTGCTGATTAACAGATGCGATG 6408
QY 2271 ATCCAGACCAATCAACATGGCACCTGCTACCTGTACAGTCAAGGATGGTAGAA 2330
Db 6407 ATCCAGACCAATCAACATGGCACCTGCTACCTGTACAGTCAAGGATGGTAGAA 6348
QY 2331 TGTCTGCTGCTCTGACACAGCAATATTACGCCATTTGCTGCTGATTAACAGAGCTCTTC 2390
Db 6347 TGTCTGCTGCTCTGACACAGCAATATTACGCCATTTGCTGCTGATTAACAGAGCTCTTC 6288
QY 2391 TAGCATAGGGCACAATTCGATCGTGGAAAGTTTGGCTTCTACCGATTTAGCAGTTTG 2450
Db 6287 TAGCATAGGGCACAATTCGATCGTGGAAAGTTTGGCTTCTACCGATTTAGCAGTTTG 6228
QY 2451 ATACACTTTTCTCTAAGTATCCACCTGAATCATATAATCGGCAAAATAGAGAAAATTTGACC 2510
Db 6227 ATACACTTTTCTCTAAGTATCCACCTGAATCATATAATCGGCAAAATAGAGAAAATTTGACC 6168
QY 2511 ATGTGTAGCGGCGCAATCTGATTCACCTGAGATGATTAATCTAGTAGAATCTCTTCGCT 2570
Db 6167 ATGTGTAGCGGCGCAATCTGATTCACCTGAGATGATTAATCTAGTAGAATCTCTTCGCT 6108
QY 2571 ATCAAAATTTCACTTCCACCTTCCACTCACCGGTTCTCCATTCATGGCTGAACCTGCTTC 2630
Db 6107 ATCAAAATTTCACTTCCACCTTCCACTCACCGGTTCTCCATTCATGGCTGAACCTGCTTC 6048
QY 2631 CTCTGTGATGACACACATCATCTCAATATCCGAATAGGCGCCATCAGTCTGACGACC 2690
Db 6047 CTCTGTGATGACACACATCATCTCAATATCCGAATAGGCGCCATCAGTCTGACGACC 5988
QY 2691 AGAGAGCCATAAACCAGCAATAGCTTAACATATCCCATATTTATCCATATTTCCGTTTC 2750
Db 5987 AGAGAGCCATAAACCAGCAATAGCTTAACATATCCCATATTTATCCATATTTCCGTTTC 5928
QY 2751 CTTAATTTTCAAGCAAACTTCT 2810
Db 5927 CTTAATTTTCAAGCAAACTTCT 5868
QY 2811 TATTCTCATCT 2870
Db 5867 TATTCTCATCT 5808
QY 2871 AGCACCGTTCTTATTCTAGCTATTATAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2930
Db 5807 AGCACCGTTCTTATTCTAGCTATTATAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5748
QY 2931 ATAAATTAATAGCATCTTAATCTTCAAACTGGCCGCTTTGTTGAATCTCTCTCTCTCTCT 2990
Db 5747 ATAAATTAATAGCATCTTAATCTTCAAACTGGCCGCTTTGTTGAATCTCTCTCTCTCTCT 5688
QY 2991 AAAATAATTTTTCGTTTCCCAATTCACATTCGAATATAGAAAAATCCATCTTCATCGGC 3050
Db 5687 AAAATAATTTTTCGTTTCCCAATTCACATTCGAATATAGAAAAATCCATCTTCATCGGC 5628
QY 3051 TTTTCTGTCATCATCTGTTATGAATCAAAATCGCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3110

Db 5627 TTTTTCGTGTCATCATCTGTATGAATCAAAATCGCTTCTTCTGTGTCATCAAGGTTTAATTT 5568
Qy 3111 TTTATGTAATTTCTTTTACAAACACACATAGAGATTAACTTTTACGGTGTAAACCTTC 3170
Db 5567 TTTATGTAATTTCTTTTACAAACACACATAGAGATTAACTTTTACGGTGTAAACCTTC 5508
Qy 3171 CTCCAATCAGACAAACGTTTCAAAATCTTTTCTTCATCATCGGTCAATAAAATCCGTATC 3230
Db 5507 CTCCAATCAGACAAACGTTTCAAAATCTTTTCTTCATCATCGGTCAATAAAATCCGTATC 5448
Qy 3231 CTTTACAGATATTTTTCAGATTTCTGCAATTCGCGATTGTATCCGATTTATATTTATT 3290
Db 5447 CTTTACAGATATTTTTCAGATTTCTGCAATTCGCGATTGTATCCGATTTATATTTATT 5388
Qy 3291 TTTTCGGTATTTTATTATAA 3310
Db 5387 TTTTCGGTCAATCAATTTGAA 5368

RESULT 2
US-08-398-028B-1/c
; Sequence 1, Application US/08398028B
; Patent No. 5780285
; GENERAL INFORMATION:
; APPLICANT: Marcus D. Ballinger and James A. Wells
; TITLE OF INVENTION: SUBSTITUTED VARIANTS CAPABLE OF CLEAVING
; TITLE OF INVENTION: SUBSTRATES CONTAINING DIBASIC RESIDUES
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/398,028B
; FILING DATE: 03-Mar-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kubinec, Jeffrey S.
; REGISTRATION NUMBER: 36,575
; REFERENCE/DOCKET NUMBER: P0936
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8228
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8119 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-08-398-028B-1

Query Match 48.0%; Score 2467.2; DB 1; Length 8119;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2472; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 831 TTGAGAAACCGAGCTGTAAAGTACAGTGGCAATTTATCTCATATTATAAAGCCAGTCAAT 890
Db 7847 TTGAGAAACCGAGCTGTAAAGTACAGTGGCAATTTATCTCATATTATAAAGCCAGTCAAT 7788
Qy 891 TAGGCCTATCTGCAATTTCTGTAATAGATTCAAAATCTCTGATCAATCAATCAAC 950
Db 7787 TAGGCCTATCTGCAATTTCTGTAATAGATTCAAAATCTCTGATCAATCAATCAAC 7728
Qy 951 AAACAGATGATGCTCTGTAAAGTACAGTGGCAATTTATCTCATATTATAAAGCCAGTCAAT 1010

Db 7727 AAACAGATGATGCTCTGTAAAGTACAGTGGCAATTTATCTCATATTATAAAGCCAGTCAAT 7668
Qy 1011 ATTTTCTCTGTAATAATCGGTAGAGGTAATTAATTAATTTATTTATTTATTTAAGTTAA 1070
Db 7667 ATTTTCTCTGTAATAATCGGTAGAGGTAATTAATTAATTTATTTATTTAAGTTAA 7608
Qy 1071 ACCCAGTAATGAAGTCCATGTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1130
Db 7607 ACCCAGTAATGAAGTCCATGTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7548
Qy 1131 TTTTGGGAAACAAATTTTCCCGAACCAATTAATTTCTCTACATCAGAGAGAGAGAGAGAGAGAG 1190
Db 7547 TTTTGGGAAACAAATTTTCCCGAACCAATTAATTTCTCTACATCAGAGAGAGAGAGAGAGAG 7488
Qy 1191 AAAATCTTTTGAAGTCAATTTTACAGGAGTCCAAATACAGAGAGAGAGAGAGAGAGAGAGAG 1250
Db 7487 AAAATCTTTTGAAGTCAATTTTACAGGAGTCCAAATACAGAGAGAGAGAGAGAGAGAGAGAG 7428
Qy 1251 CATCAAAATTTGTATAAAGTGGCTTAACCTTATCCCAATAACCTAACTCTCCGTCGCTAT 1310
Db 7427 CATCAAAATTTGTATAAAGTGGCTTAACCTTATCCCAATAACCTAACTCTCCGTCGCTAT 7368
Qy 1311 TGTAACCAAGTCTAAAGAGTGTATTTGAGTGTATCAACCTTGTCTCACTAGAGAGAGAGAGAG 1370
Db 7367 TGTAACCAAGTCTAAAGAGTGTATTTGAGTGTATCAACCTTGTCTCACTAGAGAGAGAGAGAG 7308
Qy 1371 CAGGTAATAATTTATATCT 1430
Db 7307 CAGGTAATAATTTATATCT 7248
Qy 1431 CTGTGTTATACATAAAGTCTGTTTGTGTTTCAAAATGAATTAATTAATTAATTAATTAATTAAT 1490
Db 7247 CTGTGTTATACATAAAGTCTGTTTGTGTTTCAAAATGAATTAATTAATTAATTAATTAATTAAT 7188
Qy 1491 TCCAAATTTCTAAATCAATTTTATTAAGTGTATTTGATATGCTCTCTAAATTTTATCT 1550
Db 7187 TCCAAATTTCTAAATCAATTTTATTAAGTGTATTTGATATGCTCTCTAAATTTTATCT 7128
Qy 1551 AAAGTGAATTTAGGAGGCTTACTTGTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1610
Db 7127 AAAGTGAATTTAGGAGGCTTACTTGTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 7068
Qy 1611 AGTCAATATTACTGTATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1670
Db 7067 AGTCAATATTACTGTATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 7008
Qy 1671 GTTTGTTGAATTAATCGGTGCTTTAGTTGAAGATTAAGACCAATTAAGACCAATTAAGACCAATTAAG 1730
Db 7007 GTTTGTTGAATTAATCGGTGCTTTAGTTGAAGATTAAGACCAATTAAGACCAATTAAGACCAATTAAG 6948
Qy 1731 TTTTGTGCTTTTCT 1790
Db 6947 TTTTGTGCTTTTCT 6888
Qy 1791 AATAAGGTAATTTGCGGGATAGACTGTACATCTCTCAGCATATAAATCCCTTTCTCA 1850
Db 6887 AATAAGGTAATTTGCGGGATAGACTGTACATCTCTCAGCATATAAATCCCTTTCTCA 6828
Qy 1851 TTTTCTTAATGTAATCTATTACCTTATTATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1910
Db 6827 TTTTCTTAATGTAATCTATTACCTTATTATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6768
Qy 1911 TCTTATTACGAAATAGGCGGATTTAAGCACACCCCTTTTATTCGGTTAATTCGGCCATGAC 1970
Db 6767 TCTTATTACGAAATAGGCGGATTTAAGCACACCCCTTTTATTCGGTTAATTCGGCCATGAC 6708
Qy 1971 AGCCATGATATTACTATACTAGAGAGAGTTAATAAATACGTAACCAACATGATTAAACA 2030
Db 6707 AGCCATGATATTACTATACTAGAGAGAGTTAATAAATACGTAACCAACATGATTAAACA 6648
Qy 2031 ATTATTAGAGTCTATCGTTCAAAATGGTATCGGTTTGTGACACATCCCATATATATTCGGTG 2090

6647 ATTATTAGAGGTCATCGTTCAAAATGGTATGCGTTTGGACACATCCACTATATATCCGGT 6588
2091 TGGTTCTGTCACCTCCTGAATCCCATTCAGAAATCTCTAGCGATTCAGAGTTTCTC 2150
6587 TGGTTCTGTCACCTCCTGAATCCCATTCAGAAATCTCTAGCGATTCAGAGTTTCTC 6528
2151 AGAGTCGGAAGTGGACAGACATTAACGAATGGCAGAGATGGTCAATAACCTGGAAGG 2210
6527 AGAGTCGGAAGTGGACAGACATTAACGAATGGCAGAGATGGTCAATAACCTGGAAGG 6468
2211 ATCTGATTCGTTTACCTGTCAGTTAGACCGAGCGCTCTGTATTAACAGATGCGATG 2270
6467 ATCTGATTCGTTTACCTGTCAGTTAGACCGAGCGCTCTGTATTAACAGATGCGATG 6408
2271 ATGACAGCAATCAACATGGACCTGCCATTCCTACCTGTACAGTCAAGGATGGTAGAAA 2330
6407 ATGACAGCAATCAACATGGACCTGCCATTCCTACCTGTACAGTCAAGGATGGTAGAAA 6348
2331 TGTTCGCGTCTTGCACACGAATATTAGGCAATTTGCTGCTGATTAACAGCTCTTC 2390
6347 TGTTCGCGTCTTGCACACGAATATTAGGCAATTTGCTGCTGATTAACAGCTCTTC 6288
2391 TAGGATAGGACAAATCGCATCTGGAAGTTTGGCTTCTACCGATTTAGCAGTTG 2450
6287 TAGGATAGGACAAATCGCATCTGGAAGTTTGGCTTCTACCGATTTAGCAGTTG 6228
2451 ATACACTTTCTCTAAGTATCCACTGAATCATATAATCGGCAAAATAGAGAAAAATTGACC 2510
6227 ATACACTTTCTCTAAGTATCCACTGAATCATATAATCGGCAAAATAGAGAAAAATTGACC 6168
2511 ATGTTAGCGGCAATCTGATTCACCTGAGATGATTAATCTAGTAGATCTCTTCGCT 2570
6167 ATGTTAGCGGCAATCTGATTCACCTGAGATGATTAATCTAGTAGATCTCTTCGCT 6108
2571 ATCAAAATTCATCTCCACTTCCACTCCCGGTTGTCCATTCATCGCTGAACTCTGCTTC 2630
6107 ATCAAAATTCATCTCCACTTCCACTCCCGGTTGTCCATTCATCGCTGAACTCTGCTTC 6048
2631 CTCTGTTGACATGACACATCATCTCAATTCGAAATAGCGGCCCATCAGTCTGACGACC 2690
6047 CTCTGTTGACATGACACATCATCTCAATTCGAAATAGCGGCCCATCAGTCTGACGACC 5988
2691 AAGAGAGCCATAAACAAGATAGCTTAACTATCCCATTCATTCATTCATTCATTCATTC 2750
5987 AAGAGAGCCATAAACAAGATAGCTTAACTATCCCATTCATTCATTCATTCATTCATTC 5928
2751 CTAAATTCATGAACAATCT 2810
5927 CTAAATTCATGAACAATCT 5868
2811 TATTCTCATCT 2870
5867 TATTCTCATCT 5808
2871 AGCAGCGTTCTTATTAGCATTAATACTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2930
5807 AGCAGCGTTCTTATTAGCATTAATACTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5748
2931 ATAACAATATAGCATTAATACT 2990
5747 ATAACAATATAGCATTAATACT 5688
2991 AAAATAAATTTTCCGTTCCCAATTCACATTCGCAATTAAGAAAAATCCATCTTCATCGGC 3050
5687 AAAATAAATTTTCCGTTCCCAATTCACATTCGCAATTAAGAAAAATCCATCTTCATCGGC 5628
3051 TTTTTCGTCATCATCTGTATGAATCAATCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3110
5627 TTTTTCGTCATCATCTGTATGAATCAATCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5568
3111 TTTTATGATTTCTTTTAAACAAACCAACCATAGGAGATTAACTTTTACGGTGTAAACCTTC 3170
5567 TTTTATGATTTCTTTTAAACAAACCAACCATAGGAGATTAACTTTTACGGTGTAAACCTTC 5508

3171 CTCGAATCAGACAAACGTTTCAAAATCTTTTCTTCATCATCGGTCAATAAAATCGGTATC 3230
5507 CTCGAATCAGACAAACGTTTCAAAATCTTTTCTTCATCATCGGTCAATAAAATCGGTATC 5448
3231 CTTTACAGATATTTTGCAGTTTCCTCAATTCGCCGATTCGATATCCGATTTATTTATTTATT 3290
5447 CTTTACAGATATTTTGCAGTTTCCTCAATTCGCCGATTCGATATCCGATTTATTTATTTATT 5388
3291 TTTTCGATTTTATTTTAA 3310
5387 TTTTCGATTCGAATTTGAA 5368

RESULT 3
US-08-504-265B-1/c
; Sequence 1, Application US/08504265B
; Patent No. 5837516
; GENERAL INFORMATION:
; APPLICANT: Marcus D. Ballinger and James A. Wells
; TITLE OF INVENTION: SUBSTITUTION VARIANTS CAPABLE OF CLEAVING
; TITLE OF INVENTION: SUBSTRATES CONTAINING BASIC RESIDUES
; NUMBER OF SEQUENCES: 90
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/504,265B
; FILING DATE: 19-Jul-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/398028
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kubinec, Jeffrey S.
; REGISTRATION NUMBER: 36,575
; REFERENCE/DOCKET NUMBER: P0936P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8228
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8119 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-08-504-265B-1

Query Match 48.0%; Score 2467.2; DB 2; Length 8119;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2472; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
831 TTAGAAAAACCGACTGTAAAGATGACAGTCGGGATTCATCTATTTATTAAGCCAGTCAT 890
7847 TTAGAAAAACCGACTGTAAAGATGACAGTCGGGATTCATCTATTTATTAAGCCAGTCAT 7788
891 TAGGCTATCTGACAAATCTCTGAATAGAGTTTCATAAAACAAATCTCTGATGAACCATCAC 950
7787 TAGGCTATCTGACAAATCTCTGAATAGAGTTTCATAAAACAAATCTCTGATGAACCATCAC 7728
951 AAACAGAAATGATGATCTGATGAAGATAGCGGTAAATATATTTGAATACCTTTTATTAATGA 1010
7727 AAACAGAAATGATGATCTGATGAAGATAGCGGTAAATATATTTGAATACCTTTTATTAATGA 7668

Db 5507 CTCCAATCAGACAAACGTTTCAAAATCTTTTCTTCATCGGTGCATATAAAATCCGTATC 5448
Qy 3231 CTTTACAGATATTTGAGTTTTCGTCATTTGCGATTCGCGATTCATATCCGATTTATATTTAT 3290
Db 5447 CTTTACAGATATTTGAGTTTTCGTCATTTGCGATTCGCGATTCATATCCGATTTATATTTAT 5388
Qy 3291 TTTTCGTTATTTTATTTAAA 3310
Db 5387 TTTTCGTTCAATCATTTGGA 5368

RESULT 4
US-08-992-334-1
; Sequence 1, Application US/08992334
; Patent No. 5919678
; GENERAL INFORMATION:
; APPLICANT: Gruss, Alexandra
; APPLICANT: Maguin, Emmanuelle
; TITLE OF INVENTION: METHODS FOR USING A TEMPERATURE-SENSITIVE
; TITLE OF INVENTION: PLASMID
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Christie Parker & Hale, LLP
; STREET: 350 West Colorado Boulevard, Suite 500
; CITY: Pasadena
; STATE: California
; COUNTRY: United States
; ZIP: 91105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/992,334
; FILING DATE: 17-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/302,752
; FILING DATE: 24-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00248
; FILING DATE: 12-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 992/03034
; FILING DATE: 13-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Prout, D. Bruce
; REGISTRATION NUMBER: 20958
; REFERENCE/DOCKET NUMBER: C93:31779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (626) 795-9900
; TELEFAX: (626) 577-8800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3792 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: YES
; IMMEDIATE SOURCE:
; CLONE: pg-host4
; US-08-992-334-1

Query Match 31.8%; Score 1636.8; DB 2; Length 3792;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 1686; Conservative 0; Mismatches 27; Indels 3; Gaps 3;
Qy 3279 TTTATATTTATTTTTCGTTATTTTATTTAAACGTTCTCAAAATCGTTTCTCGGACGTTT 3338
Db 1181 TATATATTTATTTATCGCATTTTATTTAAACGTTCTCAAAATCGTTTCTCGGACGTTT 1240

Qy 3339 TAGCGTTTATTTTCGTTTATTCGGCATATTCGTTTAAACAGCGTTTATCGTAGCGTAA 3398
Db 1241 TAGCGTTTATTTTCGTTTATTCGGCATATTCGTTTAAACAGCGTTTATCGTAGCGTAA 1300
Qy 3399 AAGCCCTTGAGCGTAGCGT-GCTTTGACGGAAGATGTTGTTCTGTATGATATTAAGAGCC 3457
Db 1301 AAGCCCTTGAGCGTAGCGTGTGTTGACGCGAAGATGTTGTTCTGTATGATATTAAGAGCC 1360
Qy 3458 GATGACTGAATGAAATAATAAGCGCAGCGTCTCTATTTTCGTTGAGGAGGCTCAAGG 3517
Db 1361 GATGACTGAATGAAATAATAAGCGCAGCGCTTCTATTTTCGTTGAGGAGGCTCAAGG 1420
Qy 3518 GAGTTTGAGGGAATGAAATTCCTCATCGGTTTGATTTTAAATAATGCTTGCATTTTTCG 3577
Db 1421 GAGTATGAGGGAATGAAATTCCTCATCGGTTTGATTTTAAATAATGCTTGCATTTTTCG 1480
Qy 3578 CGAGCGGTAGCGCTGAAATTTTGAATAAATTTGGAATTTGGAATAAATGCGGGA 3637
Db 1481 CGAGCGGTAGCGCTGAAATTTTGAATAAATTTTGAATAAATTTGGAATAAATGCGGGA 1540
Qy 3638 AAGGAAGCGAATTTTTCGTTTCCGTACTACGACCCCATTAAGTCGCGAGTGCCTTTT 3697
Db 1541 AAGGAAGCGAATTTTTCGTTTCCGTACTACGACCCCATTAAGTCGCGAGTGCCTTTT 1600
Qy 3698 GTGCCAAAACGCTCTATCCCACTGGCTCAAGGGTTTTCGAGGGTTTTCATCGCCCAAC 3757
Db 1601 GTGCCAAAACGCTCTATCCCACTGGCTCAAGGGTTTTCGAGGGTTTTCATCGCCCAAC 1660
Qy 3758 GAATCGCCAAACGTTTTCGCAACGTTTCTATAATCTATAATTAAGTAGCTTTATGTT 3817
Db 1661 GAATCGCCAAACGTTTTCGCAACGTTTCTATAATCTATAATTAAGTAGCTTTATGTT 1720
Qy 3818 GTTTTATGATTAACAAGTATACACTATTTTATAAATAATTTGATTTGAGAGTTTATA 3877
Db 1721 GTTTTATGATTAACAAGTATACACTATTTTATAAATAATTTGATTTGAGAGTTTATA 1780
Qy 3878 AATGTTGATTTTCAGAAATCGAAAAAAGAGTTATGATTTCTCTGACAAAAGACAGATAA 3937
Db 1781 AATGTTGATTTTCAGAAATCGAAAAAAGAGTTATGATTTCTCTGACAAAAGACAGATAA 1840
Qy 3938 AAAATTAACAGATATCGCGAACAAGAGTTTTCAGAAATCTCGGTTTCGCGGTATGAC 3997
Db 1841 AAAATTAACAGATATCGCGAACAAGAGTTTTCAGAAATCTCGGTTTCGCGGTATGAC 1900
Qy 3998 TATAGAAGATATGCAAGAAAGGAATCAGAAACAAAAAATAAGCGAAAGCTCGCGTTT 4057
Db 1901 TATAGAAGATATGCAAGAAAGGAATCAGAAACAAAAAATAAGCGAAAGCTCGCGTTT 1960
Qy 4058 TAGAAGGATACGAGTTTTCGTTACTTTGTTGTTTGAAGGTA-TATATCATGGCTATTAA 4116
Db 1961 TAGAAGGATACGAGTTTTCGTTACTTTGTTGTTTGAAGGTA-TATATCATGGCTATTAA 2020
Qy 4117 AATATCTTAAGCTAGAAATTTTCGATTTTATATATCTGACTCAATTCCTAATGATTG 4176
Db 2021 AATATCTTAAGCTAGAAATTTTCGATTTTATATATCTGACTCAATTCCTAATGATTG 2080
Qy 4177 GAAAGAAAAATTAGAGAGTTTGGCGGTATCTATGCGTGTGAGTCTTTTACGATATGA 4236
Db 2081 GAAAGAAAAATTAGAGAGTTTGGCGGTATCTATGCGTGTGAGTCTTTTACGATATGA 2140
Qy 4237 CGAAAAAAGATAGATACATGGAATAGTACTGATTTATACGAAATGGAAGACATA 4296
Db 2141 CGAAAAAAGATAGATACATGGAATAGTACTGATTTATACAAAAATGGAAGACATA 2200
Qy 4297 TAAAAAACCACTATCACGTTATATATTTGACGAAATCTGTAAACATAGAAAGCGT 4356
Db 2201 TAAAAAACCACTATCACGTTATATATTTGACGAAATCTGTAAACATAGAAAGCGT 2260
Qy 4357 TAGGAACAAAGATTAAGGGAATTTGGGATAGTTCAGTTGCTCATGTTGAGATATCTGA 4416
Db 2261 TAGGAACAAAGATTAAGGGAATTTGGGATAGTTCAGTTGCTCATGTTGAGATATCTGA 2320

QY 4417 TTATATCAAGGTTTCATATGAATATTTGACTCATGAATCAAAAGGACGCTATTGCTAAGAA 4476
Db 2321 TTATATCAAGGTTTCATATGAATATTTGACTCATGAATCAAAAGGACGCTATTGCTAAGAA 2380
QY 4477 TAAACATATATACGCAAAAGAGATATTTGAAACATTAATGATTTTGATATTTGACCGCTA 4536
Db 2381 TAAACATATATACGCAAAAGAGATATTTGAAACATTAATGATTTTGATATTTGACCGCTA 2440
QY 4537 TATAACACTTGATGAAGCAAAAGAGATATTTGAAACATTAATGATTTTGATATTTGACCGCTA 4596
Db 2441 TATAACACTTGATGAAGCAAAAGAGATATTTGAAACATTAATGATTTTGATATTTGACCGCTA 2500
QY 4597 TGACTATATTTGTTGAATACAAAGATTTAATGCTTTTATTCGCCCTTAGCGGAGCGGA 4656
Db 2501 TGACTATATTTGTTGAATACAAAGATTTAATGCTTTTATTCGCCCTTAGCGGAGCGGA 2560
QY 4657 GTTTCGAATTTTAAATACGAATGATGTAAGAGATATTTGTTTCAACAAACTCTAGCGCCTT 4716
Db 2561 GTTTCGAATTTTAAATACGAATGATGTAAGAGATATTTGTTTCAACAAACTCTAGCGCCTT 2620
QY 4717 TAGATTATGTTTGGGCAATATATCAGTGTGATATAGCAAGGTTATGCAAGGTTCT 4776
Db 2621 TAGATTATGTTTGGGCAATATATCAGTGTGATATAGCAAGGTTATGCAAGGTTCT 2680
QY 4777 TGATCTGAAACGGGGGAAATTAATATGACAAACAAAGAGAGATATTTGCTGAAAT 4836
Db 2681 TGATCTGAAACGGGGGAAATTAATATGACAAACAAAGAGAGATATTTGCTGAAAT 2740
QY 4837 GAGGAATTAATAAGAAATTAAGGACTTAAAGAGCGTATTGAAAGATACAGAGAAATG 4896
Db 2741 GAGGAATTAATAAGAAATTAAGGACTTAAAGAGCGTATTGAAAGATACAGAGAAATG 2800
QY 4897 GAAGTTGAATTAAGTACAAATAGATTTATTGAGGAGGAGGATTTGAAATTAATAA 4956
Db 2801 GAAGTTGAATTAAGTACAAATAGATTTATTGAGGAGGAGGATTTGAAATTAATAA 2860
QY 4957 GCGCCCTGACGAAGTCCGACTTCGTTCTTTTT 4992
Db 2861 G-CGCCCTGACGAAGTCCGAGGGGTTTTTATTTT 2895

RESULT 5

US-08-302-752-1
; Sequence 1, Application US/08302752
; Patent No. 6025190
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: THERMOSENSIBLE PLASMID
; NUMBER OF SEQUENCES: 3
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/302,752
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9203034
; FILING DATE: 13-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO FR/93/00248
; FILING DATE: 12-MAR-1993
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3792 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

US-08-302-752-1

Query Match 31.9%; Score 1636.8; DB 3; Length 3792;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 1686; Conservative 0; Mismatches 27; Indels 3; Gaps 3;
QY 3279 TTTATATTTATTTTTCGGTATTTTATTAATAAAGCTCTCAAAATCGTTTCTCGGACGTTT 3338
Db 1181 TATATATATTTATTCGCATTTTTTATTAATAAAGCTCTCAAAATCGTTTCTCGAGCGTTT 1240
QY 3339 TAGCGTTTATTTTCGTTAGTTATTCGCATTAATCGTTTAAACAGCGCTATCGTAGCGTAA 3398
Db 1241 TAGCGTTTATTTTCGTTAGTTATTCGCATTAATCGTTTAAACAGCGCTATCGTAGCGTAA 1300
QY 3399 AAGCCCTTGAGCGTAGCGT-GCTTTCAGCGAGAGATGTTGCTGTAGATATTAAGAGCC 3457
Db 1301 AAGCCCTTGAGCGTAGCGTGGCTTTCAGCGAGAGATGTTGCTGTAGATATTAAGAGCC 1360
QY 3458 GATGACTGAATGAATAAATAAGCGCAGCGTCTTCTTATTTTCGTTTCGAGAGGCTCAAGG 3517
Db 1361 GATGACTGAATGAATAAATAAGCGCAGCGCTCTTATTTTCGTTTCGAGAGGCTCAAGG 1420
QY 3518 GAGTTTGAGGGAATGAATTCCTCATGCTGATTTTAAATTTGAAATTTGCAATTTTGC 3577
Db 1421 GAGTTTGAGGGAATGAATTCCTCATGCTGATTTTAAATTTGCAATTTTGC 1480
QY 3578 CGAGCGGTAGCGCTGGAATAATTTTGAATAAATTTGAAATTTGAAATTTGCGGGGA 3637
Db 1481 CGAGCGGTAGCGCTGGAATAATTTTGAATAAATTTGAAATTTGCGGGGA 1540
QY 3638 AAGAGCGGAATTTTCGTTCCGCTACTAGACCCCGCATTAAGTCCGAGTGCCAAATTTT 3697
Db 1541 AAGAGCGGAATTTTCGTTCCGCTACTAGACCCCGCATTAAGTCCGAGTGCCAAATTTT 1600
QY 3698 GTGCCAAAACGCTCTATCCAACTGGCTCAAGGGTTTGAGGGGTTTTTCAATCGCGAAC 3757
Db 1601 GTGCCAAAACGCTCTATCCAACTGGCTCAAGGGTTTGAGGGGTTTTTCAATCGCGAAC 1660
QY 3758 GAATCGCCAAACGTTTTGCGCAACGTTTTTATTAATCTATTTAAGTAGCTTTATGTT 3817
Db 1661 GAATCGCCAAACGTTTTGCGCAACGTTTTTATTAATCTATTTAAGTAGCTTTATGTT 1720
QY 3818 GTTTTATGATTACAAAGGATACACTAATTTTATAAATTTATGATTTGAGGTTTTT 3877
Db 1721 GTTTTATGATTACAAAGGATACACTAATTTTATAAATTTATGATTTGAGGTTTTT 1780
QY 3878 AATGGTGATTTTCAAGATTCGAAAAAAGAGTATGATTTCTCTGACAAAAAGAGCAAGATA 3937
Db 1781 AATGGTGATTTTCAAGATTCGAAAAAAGAGTATGATTTCTCTGACAAAAAGAGCAAGATA 1840
QY 3938 AAAATTAACAGATATGCGGCAACAAAGGTTTTTCAAAATCTCGCGTTGCGCGCTAGC 3997
Db 1841 AAAATTAACAGATATGCGGCAACAAAGGTTTTTCAAAATCTCGCGTTGCGCGCTAGC 1900
QY 3998 TATAGAAGATATGCAAGAAAGGAATCAGAACAAAAAATAAGCAAGCTCGCGTTTT 4057
Db 1901 TATAGAAGATATGCAAGAAAGGAATCAGAACAAAAAATAAGCAAGCTCGCGTTTT 1960
QY 4058 TAGAAGATACAGGTTTTTCGCTACTTTTGTGATTAAGGTAA-TATATCATGCTATTAA 4116
Db 1961 TAGAAGATACAGGTTTTTCGCTACTTTTGTGATTAAGGTAA-TATATCATGCTATTAA 2020
QY 4117 AAATACCTAAAGCTAGAAAAATTTTCGATTTTATTTATATCTGACTCAATTCCTAATGATG 4176
Db 2021 AAATACCTAAAGCTAGAAAAATTTTCGATTTTATTTATATCTGACTCAATTCCTAATGATG 2080
QY 4177 GAAAGAAAAATTAGAGAGTTTGGCGTATCTATGCTGTGCTGAGTCCCTTTACACGATATGA 4236
Db 2081 GAAAGAAAAATTAGAGAGTTTGGCGTATCTATGCTGTGCTGAGTCCCTTTACACGATATGA 2140
QY 4237 CGAAAAAAGATAAAGATACATGGAATAGTAGTGTGTTATACGAATTTGAAAGCACTA 4296
Db 2141 CGAAAAAAGATAAAGATACATGGAATAGTAGTGTGTTATATACAAATTTGAAAGCACTA 2200
QY 4297 TAAAAAACCACTATCACGTTTATATATATTTGACCAAGAAATCCCTGTAAACAATAGAAAGCGT 4356

3818 GTTTTATGATTACAAAGTGATACACTAAATTTTATAAATTTATTTGATTTGGAGTTTTTA 3877
Db |||||
3163 GTTTTATGATTACAAAGTGATACACTAAATTTTATAAATTTATTTGATTTGGAGTTTTTA 3222
Qy |||||
3878 AATGTTGATTTGAGAAATCGAAAAAGAGTTATGATTTCTCTGACAAAGAGCAAGATAA 3937
Db |||||
3223 AATGTTGATTTGAGAAATCGAAAAAGAGTTATGATTTCTCTGACAAAGAGCAAGATAA 3282
Qy |||||
3938 AAAATTTACAGATATGCGGAAACAAAGAGTTTTCAAAAATCTCGGTTGCGGCTTAGC 3997
Db |||||
3283 AAAATTTACAGATATGCGGAAACAAAGAGTTTTCAAAAATCTCGGTTGCGGCTTAGC 3342
Qy |||||
3998 TATAGAAGATATGCAAGAAAGGATTCAGAACAAAAAATAGCGAAAGTCGCGTTTT 4057
Db |||||
3343 TATAGAAGATATGCAAGAAAGGATTCAGAACAAAAAATAGCGAAAGTCGCGTTTT 3402
Qy |||||
4058 TAGAAGGATACAGATTTTTCGCTACTGTTTGTGTAAGGTAA-TATATCATGGCTATTAA 4116
Db |||||
3403 TAGAAGGATACAGATTTTTCGCTACTGTTTGTGTAAGGTAA-TATATCATGGCTATTAA 3462
Qy |||||
4117 AAATCTAAAGCTAGAAATTTTGGATTTTATTTATATCTCTGACTCAATTCCTAATGTTG 4176
Db |||||
3463 AAATCTAAAGCTAGAAATTTTGGATTTTATTTATCTCTGACTCAATTCCTAATGTTG 3522
Qy |||||
4177 GAAAGAAAATTTAGAGATTTGGGCGTATCTATGCTGTCTAGTCTTTTACAGATATGGA 4236
Db |||||
3523 GAAAGAAAATTTAGAGATTTGGGCGTATCTATGCTGTCTAGTCTTTTACAGATATGGA 3582
Qy |||||
4237 CGAAAAAAGGATAAGATACATGAAATAGTAGTGTATGATGATATACGAAATGGAAGCACTA 4296
Db |||||
3583 CGAAAAAAGGATAAGATACATGAAATAGTAGTGTATGATGATATACGAAATGGAAGCACTA 3642
Qy |||||
4297 TAAAAACCACTATACGTTATATATATGTCGACGAAATCTGTAAACATAGAAAGCGT 4356
Db |||||
3643 TAAAAACCACTATACGTTATATATATGTCGACGAAATCTGTAAACATAGAAAGCGT 3702
Qy |||||
4357 TAGGACACAGATTTAGCGAAATTTGGGAATAGTTTCACTGTCTATGTTGAGATCTTGA 4416
Db |||||
3703 TAGGACACAGATTTAGCGAAATTTGGGAATAGTTTCACTGTCTATGTTGAGATCTTGA 3762
Qy |||||
4417 TTATATCAAAGGTTTCATATGAATTTTGACTCATGAATCAAAGGACGCTATTGCTAAGAA 4476
Db |||||
3763 TTATATCAAAGGTTTCATATGAATTTTGACTCATGAATCAAAGGACGCTATTGCTAAGAA 3822
Qy |||||
4477 TAAACATATATACGACAAAAGATATTTTGAACATTAATGATTTTGTATTTGACCGCTA 4536
Db |||||
3823 TAAACATATATACGACAAAAGATATTTTGAACATTAATGATTTTGTATTTGACCGCTA 3882
Qy |||||
4537 TAAACATTTGATGAAAGCCAAAAAGAGATTTGAAGAAATTTTACTTTTAGATATAGTGA 4596
Db |||||
3883 TAAACATTTGATGAAAGCCAAAAAGAGATTTGAAGAAATTTTACTTTTAGATATAGTGA 3942
Qy |||||
4597 TGACTATAATTTGCTGAAATACAAAAGATTTTATGCTTTTATTCGCTTAGGGAGCGGA 4656
Db |||||
3943 TGACTATAATTTGCTGAAATACAAAAGATTTTATGCTTTTATTCGCTTAGGGAGCGGA 4002
Qy |||||
4657 GTTTGGAATTTTAAATACCAATGATGTAAGATATTTGTTCAACAAACTCTAGCGCCTT 4716
Db |||||
4003 GTTTGGAATTTTAAATACCAATGATGTAAGATATTTGTTCAACAAACTCTAGCGCCTT 4062
Qy |||||
4717 TAGATTATGTTTGGGGCAATTTATCAGTGTGATATAGAGCAAGTTATGCAAGGTTCT 4776
Db |||||
4063 TAGATTATGTTTGGGGCAATTTATCAGTGTGATATAGAGCAAGTTATGCAAGGTTCT 4122
Qy |||||
4777 TGATGCTGAAACGGGGCAATTTAAATGACAAACAAAGAAAGAGTTATTTCTGCTGAAAT 4836
Db |||||
4123 TGATGCTGAAACGGGGCAATTTAAATGACAAACAAAGAAAGAGTTATTTCTGCTGAAAT 4182
Qy |||||
4837 GAGGAATTTAAAAAAGAAATTAAGGACTTTAAAGAGCGTATTGAAAGATACAGAGAAATG 4896
Db |||||
4183 GAGGAATTTAAAAAAGAAATTAAGGACTTTAAAGAGCGTATTGAAAGATACAGAGAAATG 4242
Qy |||||
4897 GAAGTTGAATTAAGTACAAATAGATTTATTGAGAGGAGGATTTATTGAATAAATAAAA 4956

Db 4243 GAAGTTGAATTAAGTACAAATAGATTTATTGAGAGGAGGATTTATTGAATAAATAAAA 4302
Qy 4957 GCCCCTCAGCAAGAGTCGCGACTTTCGTTCTTTT 4992
Db 4303 G-CCCCCTGACCAAGAGTCGAAGGGGTTTATTATTT 4337
RESULT 8
US-08-992-334-3
; Sequence 3, Application US/08992334
; Patent No. 5919678
; GENERAL INFORMATION:
; APPLICANT: Gruss, Alexandra
; APPLICANT: Maguin, Emmanuelle
; TITLE OF INVENTION: METHODS FOR USING A TEMPERATURE-SENSITIVE
; TITLE OF INVENTION: PLASMID
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christie Parker & Hale, LLP
; STREET: 350 West Colorado Boulevard, Suite 500
; CITY: Pasadena
; STATE: California
; COUNTRY: United States
; ZIP: 91105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/992,334
; FILING DATE: 17-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/302,752
; FILING DATE: 24-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00248
; FILING DATE: 12-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/03034
; FILING DATE: 13-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Prout, D. Bruce
; REGISTRATION NUMBER: 20958
; REFERENCE/DOCKET NUMBER: C93:31779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (626) 795-9900
; TELEFAX: (626) 577-8800
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6722 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
US-08-992-334-3
Query Match 31.8%; Score 1635.2; DB 2; Length 6722;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 1685; Conservative 0; Mismatches 28; Indels 3; Gaps 3;
Qy 3279 TTATATATTTTTCGTTATTTTATTAACAGCTCTCAAAATCGTTTCTGGAGCTTT 3338
Db 4111 TATATATATTTATATCCGATTTTATTAACAGCTCTCAAAATCGTTTCTGGAGCTTT 4170
Qy 3339 TAGGTTTATTTTCGTTTATGTTATCGCATAAATCGTTTAAACAGCGCTTATCGTAGCGTAA 3398
Db 4171 TAGGTTTATTTTCGTTTATGTTATCGCATAAATCGTTTAAACAGCGCTTATCGTAGCGTAA 4230
Qy 3399 AAGCCCTTAGCGCTAGCGT-GCTTTGACGGAAGATGTTCTGTTAGATTATGAAGCC 3457

QY 3339 TAGCGTTTATTCGTTAGTTATCGGCAATATCGTTAAACAGCGGTTATCGTAGCGTAA 3398
DB 4171 TAGCGTTTATTCGTTAGTTATCGGCAATATCGTTAAACAGCGGTTATCGTAGCGTAA 4230
QY 3399 AAGCCCTTACGCTAGCGT - CTTTGGAGGAGAGATGTTGCTGTTAGATTATGAAAGCC 3457
DB 4231 AAGCCCTTACGCTAGCGTAGCGTTCGTTGAGGAGAGATGTTGCTGTTAGATTATGAAAGCC 4290
QY 3458 GATGCTGAATGAAATATATAGCGAGCGTCTCTTCTATTTTCGTTGAGGAGCGTCAAGG 3517
DB 4291 GATGCTGAATGAAATATATAGCGAGCGTCTCTTCTATTTTCGTTGAGGAGCGTCAAGG 4350
QY 3518 GAGTTTGGAGGAGTGAATTCCTTCGTTGAGGTTGATTTTAAATTCCTTGCATTTTTCG 3577
DB 4351 GAGTTTGGAGGAGTGAATTCCTTCGTTGAGGTTGATTTTAAATTCCTTGCATTTTTCG 4410
QY 3578 CGAGCGGTAGCGCTGGAATTTTGGAAAAATTTGGAATTTGGAATTTGGAATTTGGAATTTGGA 3637
DB 4411 CGAGCGGTAGCGCTGGAATTTTGGAAAAATTTGGAATTTGGAATTTGGAATTTGGAATTTGGA 4470
QY 3638 AAGGAGCGAATTTTTCGTTACGCTAGCGACCCCGCATTAAGTCCGAGTGCCTAATTTT 3697
DB 4471 AAGGAGCGAATTTTTCGTTACGCTAGCGACCCCGCATTAAGTCCGAGTGCCTAATTTT 4530
QY 3698 GTGCCAAAAACGCTCTATCCCAACTGCGTCAAGGGTTTGAAGGGTTTCAATCGCCCAAC 3757
DB 4531 GTGCCAAAAACGCTCTATCCCAACTGCGTCAAGGGTTTGAAGGGTTTCAATCGCCCAAC 4590
QY 3758 GAATCGCCAAAGTTTTCGCTTCCGCAAGTTTTCGCTTAAATCTATTTAAGTAGCTTTATGTT 3817
DB 4591 GAATCGCCAAAGTTTTCGCTTCCGCAAGTTTTCGCTTAAATCTATTTAAGTAGCTTTATGTT 4650
QY 3818 GTTTTATGATTACAAAGTGATACACTAATTTTAAATTTATTTGATTTGAGTTTTTTA 3877
DB 4651 GTTTTATGATTACAAAGTGATACACTAATTTTAAATTTATTTGATTTGAGTTTTTTA 4710
QY 3878 AATGCTGATTTTCAAGATCGAAAAAAGAGTTATGATTTCTCTGACAAAAAGCAAGATAA 3937
DB 4711 AATGCTGATTTTCAAGATCGAAAAAAGAGTTATGATTTCTCTGACAAAAAGCAAGATAA 4770
QY 3938 AAAATTAACAGATATGCGCAACAAAAAGGTTTTCAAAACTCGGTTGCGGGTTAGC 3997
DB 4771 AAAATTAACAGATATGCGCAACAAAAAGGTTTTCAAAACTCGGTTGCGGGTTAGC 4830
QY 3998 TATAGAAGATATGCAAGAAAGGAATCAGAACAAAAAATAAGCAAGAAAGTCCGCTTT 4057
DB 4831 TATAGAAGATATGCAAGAAAGGAATCAGAACAAAAAATAAGCAAGAAAGTCCGCTTT 4890
QY 4058 TAGAAGGATACAGTTTTCGCTACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 4116
DB 4891 TAGAAGGATACAGTTTTCGCTACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 4950
QY 4117 AAATACTAAAGCTAGAAAAATTTGGAATTTTATATCTCTGACTCAATTTCTTAATGATTG 4176
DB 4951 AAATACTAAAGCTAGAAAAATTTGGAATTTTATATCTCTGACTCAATTTCTTAATGATTG 5010
QY 4177 GAAAGAAAAATAGAGAGTTTGGCGGTTATCTATGCTGCTGAGTCTTTACAGATATGGA 4236
DB 5011 GAAAGAAAAATAGAGAGTTTGGCGGTTATCTATGCTGCTGAGTCTTTACAGATATGGA 5070
QY 4237 CGAAAAAAGATAAGATACATGGAATAGTAGTGTATATGGAATGGAATGGAAGCACTA 4296
DB 5071 CGAAAAAAGATAAGATACATGGAATAGTAGTGTATATGGAATGGAATGGAAGCACTA 5130
QY 4297 TAAAAAACCACTATCATCGTTTATATATTTGACGAAATCTCTGTAACAATAGAAAGCGT 4356
DB 5131 TAAAAAACCACTATCATCGTTTATATATTTGACGAAATCTCTGTAACAATAGAAAGCGT 5190
QY 4357 TAGGACAGATAGAGGAAATTTGGGAATAGTTTCAAGTTCTCTCATGTTGAGTACTTGA 4416
DB 5191 TAGGACAGATAGAGGAAATTTGGGAATAGTTTCAAGTTCTCTCATGTTGAGTACTTGA 5250

QY 4417 TTATATCAAAAGTTTCATATGAATTTTACATCATGAATCAAGGACGCTATTGCTAAGAA 4476
DB 5251 TTATATCAAAAGTTTCATATGAATTTTACATCATGAATCAAGGACGCTATTGCTAAGAA 5310
QY 4477 TAAACATATATACGACAAAAAGATTTTGTGAACATTAATGATTTGATATGACCGCTA 4536
DB 5311 TAAACATATATACGACAAAAAGATTTTGTGAACATTAATGATTTGATATGACCGCTA 5370
QY 4537 TATACACTGATGAAAGCCAAAAAGAGAAATTTGAAGAAATTTTACATTTTATGATATAGTGA 4596
DB 5371 TATACACTGATGAAAGCCAAAAAGAGAAATTTGAAGAAATTTTACATTTTATGATATAGTGA 5430
QY 4597 TGACTATATTTTGGTAAATACAAAAAGATTTTAAATGCTTTTATTCGCTTACGCTTACGCTTACGCTT 4656
DB 5431 TGACTATATTTTGGTAAATACAAAAAGATTTTAAATGCTTTTATTCGCTTACGCTTACGCTTACGCTT 5490
QY 4657 GTTTCGAATTTTAAATACGAATGATGTAAGAAATTTTCAACAACTCTAGCGCTT 4716
DB 5491 GTTTCGAATTTTAAATACGAATGATGTAAGAAATTTTCAACAACTCTAGCGCTT 5550
QY 4717 TAGATTATGCTTTGAGGGCAATTTATCAGTGTGGATATAGACCAAGTTATGCAAAAGTTCT 4776
DB 5551 TAGATTATGCTTTGAGGGCAATTTATCAGTGTGGATATAGACCAAGTTATGCAAAAGTTCT 5610
QY 4777 TGATGCTGAAACGGGGGAAATTAATGACAAACAAAGAAAGAGTTATTTGCTGAAAT 4836
DB 5611 TGATGCTGAAACGGGGGAAATTAATGACAAACAAAGAAAGAGTTATTTGCTGAAAT 5670
QY 4837 GAGGAATTAAGAAAGAAATTAAGGACTTAAAGAGCGTATTGAAAGATACAGAGAAATG 4896
DB 5671 GAGGAATTAAGAAAGAAATTAAGGACTTAAAGAGCGTATTGAAAGATACAGAGAAATG 5730
QY 4897 GAGGTTGAATTAAGTACAAATAGATTATTTAGAGAGGAGGATTTTGAATTAATATAA 4956
DB 5731 GAGGTTGAATTAAGTACAAATAGATTATTTAGAGAGGAGGATTTTGAATTAATATAA 5790
QY 4957 GCGCCCTTACGCAAGTTCGCGACTTCTGCTTTT 4992
DB 5791 G-CGCCCTTACGCAAGTTCGCGCGGTTTATTTT 5825

RESULT 10

US-08-418-085-3/c
; Sequence 3, Application US/08418085
; Patent No. 586283
; GENERAL INFORMATION:
; APPLICANT: SLIJKHUIS, HERMAN; SELTEN, GERARDUS CORNELIS
; APPLICANT: MARIA; SMAAL, ERIC BASTIAN
; TITLE OF INVENTION: PROCESS FOR OXIDATION OF STEROIDS AND
; TITLE OF INVENTION: GENETICALLY ENGINEERED CELLS USED THEREIN
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIERMAN & MUSERLIAN
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/418,085
; FILING DATE: 06-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/054,185
; FILING DATE: 26-APR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/002,608

US-09-099-011A-3/c
; Sequence 3, Application US/09099011A
; Patent No. 6,171,836
; GENERAL INFORMATION:
; APPLICANT: SLIJKHUIS, HERMAN; SELTEN,
; APPLICANT: GERARDUS CORNELIS MARIA; SMAAL,
; APPLICANT: ERIC BASTIAAN
; TITLE OF INVENTION: PROCES FOR OXIDATION OF
; TITLE OF INVENTION: STEROIDS AND GENETICALLY ENGINEERED CELLS
; TITLE OF INVENTION: USED THEREIN
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIERMAN, MUSERLIAN & LUCAS
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/099,011A
; FILING DATE: 17-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/418,085
; FILING DATE: 06-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/054,185
; FILING DATE: 26-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/002,608
; FILING DATE: 11-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/474,857
; FILING DATE: 30-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/474,798
; FILING DATE: 16-JULY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/NL89/00072
; FILING DATE: 25-SEPT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: NL89/200904.6
; FILING DATE: 06-MAY-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: NL/88/202080.3
; FILING DATE: 03-SEP-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: CHARLES A. MUSERLIAN
; REGISTRATION NUMBER: 19,583
; REFERENCE/DOCKET NUMBER: 146,1169-
; REFERENCE/DOCKET NUMBER: CON-1-DIV-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 661-8000
; TELEFAX: (212) 661-8002
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7336 BASE PAIRS
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: UNKNOWN
; FEATURE:
; OTHER INFORMATION: PLASMID pBHA-1
US-09-099-011A-3
Query Match 29.0%; Score 1493; DB 3; Length 7336;
Best Local Similarity 98.1%; Pred. No. 0;
Matches 1511; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
1778 TTCTTATCTTGATAATAAGGTAACATCTATGCGGGGTAGACTGTAAACATTCTCAGCGATA 1837

Db	5287	AGAATATTGGAGAGACCGTTCTTATTACGCTAATTAATACTCGTCTCTCTTAAGCAATCC	5228
Qy	2918	TTCAATCCTTTTAAATAACAATTAAGCATCTAAATCTCAACAAACTGGCCGTTTGTGTA	2977
Db	5227	TTCAATCCTTTTAAATAACAATTAAGCATCTAAATCTCAACAAACTGGCCGTTTGTGTA	5168
Qy	2978	ACTACTCTTTTAATAAATAATTTTTCGGTTCGCAATTCACAAATGCAATAATAGAAAAATC	3037
Db	5167	ACTACTCTTTTAATAAATAATTTTTCGGTTCGCAATTCGCAATTAATAGAAAAATC	5108
Qy	3038	CATCTTCATCGGCCTTTTTCGTCATCATCTGTATGAATCAAAATCGCCCTTCTTCTGTGTCAT	3097
Db	5107	CATCTTCATCGGCCTTTTTCGTCATCATCTGTATGAATCAAAATCGCCCTTCTTCTGTGTCAT	5048
Qy	3098	CAAGGTTAAATTTTAAATGAATTTCTTTTAAACAAACCAATAGAGATAAACCCTTTTAC	3157
Db	5047	CAAGGTTAAATTTTAAATGAATTTCTTTTAAACAAACCAATAGAGATAAACCCTTTTAC	4988
Qy	3158	GGTGTAACCTTCCTCCAAATCAGACAAAAGTTTCAAATCTTTTCTCATCATCGGTCA	3217
Db	4987	GGTGTAACCTTCCTCCAAATCAGACAAAAGTTTCAAATCTTTTCTCATCATCGGTCA	4928
Qy	3218	TAAAAATCCGATTCCTTTTACAGGATATTTTGAGATTTCGTCAAATGCGAGATGTATATCCG	3277
Db	4927	TAAAAATCCGATTCCTTTTACAGGATATTTTGAGATTTCGTCAAATGCGAGATGTATATCCG	4868
Qy	3278	ATTATATATTTATTTTTCGGTATTTTATTTTAAACAGTCTCA	3318
Db	4867	ATTATATATTTATTTTTCGGTATTTTATTTTAAACAGTCTCA	4827

RESULT 12

US-08-470-369-21/c
Sequence 21, Application US/08470369
Patent No. 623889
GENERAL INFORMATION:
APPLICANT: Dorsers, Lambertus C. J.
APPLICANT: Wagemaker, Gerard
APPLICANT: Vos, Yvonne J.
APPLICANT: Van Leen, Robert W.
TITLE OF INVENTION: MOLECULAR CLONING AND EXPRESSION OF HUMAN
TITLE OF INVENTION: IL-3
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,369
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/321,480
FILING DATE:
APPLICATION NUMBER: 07/854,297
FILING DATE: 19-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Gracey, Nancy J.
REGISTRATION NUMBER: 28,216
REFERENCE/DOCKET NUMBER: 24615-20002.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-813-5600
TELEFAX: 415-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 21:

6007	CCAGAAGTTTCTCAGAGTCGMAAAGTTGACGACACATTAACAACTGGCCACAGATGTCAT	5948
2198	AACCTGAAGGAGATCTGATTGCTTAACTGCTTCAGTTAAGACCGAAGCGCTCGTCGTAT	2257
5947	AACCTGAAGGAAGATCTGATTGCTTAACTGCTTCAGTTAAGACCGAAGCGCTCGTCGTAT	5888
2258	AACAGATGGGATGATGCAGACCAATCAACATGGCACCTGCAATTGCTACTGTACAGTCA	2317
5887	AACAGATGGCATGATGCAGACCAATCAACATGGCACCTGCAATTGCTACTGTACAGTCA	5828
2318	AGGATGGTAGAATGTTGTTCGGTTCCTTGGACACGAATATTACGCCATTTGCCCTGCATATT	2377
5827	AGGATGGTAGAATGTTGTTCGGTTCCTTGGACACGAATATTACGCCATTTGCCCTGCATATT	5768
2378	CAAAACAGCTCTTCTACGATAAAGGCGCAAAATCGCATCTGTGGAAACGTTTGGGCTTCTACCG	2437
5767	CAAAACAGCTCTTCTACGATAAAGGCGCAAAATCGCATCTGTGGAAACGTTTGGGCTTCTACCG	5708
2438	ATTTAGAGTTTGATACACHTTTCTTAAGTATCCACTGGAATCATAAATCGGCAGAAATAG	2497
5707	ATTTAGAGTTTGATACACHTTTCTTAAGTATCCACTGGAATCATAAATCGGCAGAAATAG	5648
2498	AGAAAATTGACCATGTGAAGGGGCAATCTGATTCACCTGAGATGCATAATCTAGTA	2557
5647	AGAAAATTGACCATGTGAAGGGGCAATCTGATTCACCTGAGATGCATAATCTAGTA	5588
2558	GAATCTCTTCGCTATCAAAAATTGACCTTCCACCTTCCACTCACCGGTTGTCANTCATGGC	2617
5587	GAATCTCTTCGCTATCAAAAATTGACCTTCCACCTTCCACTCACCGGTTGTCANTCATGGC	5528
2618	TGAACCTGCTCTCTGTTGACATGACACACATCATCTCAATATCCGAATAGGGCCCAT	2677
5527	TGAACCTGCTCTCTGTTGACATGACACACATCATCTCAATATCCGAATAGGGCCCAT	5468
2678	CAGTCTGACGACCAAGAGAGCCATAAACACCAATAGCCTTAACATCATGCCCATATTTAT	2737
5467	CAGTCTGACGACCAAGAGAGCCATAAACACCAATAGCCTTAACATCATGCCCATATTTAT	5408
2738	CCAAATTCGGTTCCTTAATTTCATGAAACATCTTCATCTCTCTCTAGTCATATTATTA	2797
5407	CCAAATTCGGTTCCTTAATTTCATGAAACATCTTCATCTCTCTCTAGTCATATTATTA	5348
2798	TTGGTCCATTCACTATTCTCATTCGCCCTTTTCAGATAATTTAGATTTGCTTTTCTAAATA	2857
5347	TTGGTCCATTCACTATTCTCATTCGCCCTTTTCAGATAATTTAGATTTGCTTTTCTAAATA	5288
2858	AGAAATTTGGAGAGCACCGTTCTTATCTCAGCTATTAACTCGCTCTTCTCTAGCATATTC	2917
5287	AGAAATTTGGAGAGCACCGTTCTTATCTCAGCTATTAACTCGCTCTTCTCTAGCATATTC	5228
2918	TTCAATCCTTTTAAATAACAATTAAGCATCTAACTCTTCAACAAACTGGCCGTTTGTGTA	2977
5227	TTCAATCCTTTTAAATAACAATTAAGCATCTAACTCTTCAACAAACTGGCCGTTTGTGTA	5168
2978	ACTACTCTTTTAAATAAATAATTTTTCGGTTCCCAATTTCCCATTTGCAATATAGAAAATC	3037
5167	ACTACTCTTTTAAATAAATAATTTTTCGGTTCCCAATTTCCCAATTTGCAATATAGAAAATC	5108
3038	CATCTTCATCGGCTTTTTCGTCATCATCTGTATGAATCAAAATCGCCTTCTCTCTGTGCAT	3097
5107	CATCTTCATCGGCTTTTTCGTCATCATCTGTATGAATCAAAATCGCCTTCTCTGTGCAT	5048
3098	CAAGGTTTAAATTTTATGTAATTTCTTTTAAACAAACCCATAGAGATTAACCTTTTAC	3157
5047	CAAGGTTTAAATTTTATGTAATTTCTTTTAAACAAACCCATAGAGATTAACCTTTTAC	4988
3158	GGTGTAACCTTCTCCAAATCAGACAAACGTTTCAAAATCTTTTCTCATCATCGGTCA	3217
4987	GGTGTAACCTTCTCCAAATCAGACAAACGTTTCAAAATCTTTTCTCATCATCGGTCA	4928
3218	TAAATTCGGTATCCTTTACAGGATATTTTGAGTTTTCGTCAAATTTGCCGATTTGATATCCG	3277

```

Db      4927 TAAATCCGTAATCCTTTACAGGATATTTTCGCAGTTATTCGGATTTGCGATTTATATCCG 4886
Qy      3278 ATTTATATTTATTTTTCGGTATTTTATTTAAACGTTCTCA 3318
      |||||
      |||||
Db      4867 ATTTATATTTATTTTTCGGTTCGAATCATTGGAACITTTTACA 4827
      |||||
      |||||

RESULT 14
US-08-556-978B-79
; Sequence 79, Application US/08556978B
; Patent No. 6268169
; GENERAL INFORMATION:
; APPLICANT: FAHNESTOCK, STEPHEN F.
; TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED
; TITLE OF INVENTION: SPIDER SILK ANALOGS
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; *COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD FOR WINDOWS 95
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/556,978B
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/077,600
; FILING DATE: JUNE 15, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CR-9389-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 79:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9144 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
US-08-556-978B-79

```

Query Match	29.0%; Score 1493; DB 3; Length 9144;
Best Local Similarity	98.1%; Pred. No. 0;
Matches 1511; Conservative 0; Mismatches 30; Indels 0; Gaps 0	
Qy 1778	TTCTTACTTCATATAAGGTAACATTGCGGGATAGACTGTAAACATCTTCACGCATA 1833
Db 6526	TACTCTTTTAAATATCCCGNCTGGCAATGCGGGATAGACTGTAACAATTCTCAGCATA 6588
Qy 1838	AAATCCCTTTTCATTTTCTAATGAATCTATTACCTTTATTATTAATTCAATTCGCTCAT 1893
Db 6586	AAATCCCTTTTCATTTTCTAATGAATCTATTACCTTTATTATTAATTCAAATTCGCTCAT 6645
Qy 1898	AATTAATCCCTTTTTCCTTTATTACGCCAAATGGCCGGATTTAAGCACACCCCTTTATTCCCGTT 1953
Db 6646	AATTAATCCCTTTTTCCTTTATTACGCCAAATGGCCGGATTTAAGCACACCCCTTTATTCCCGTT 6705
Qy 1958	AATGGCCCATGACGCCATGATAATTACTAATCTAGGAGAAGTTAATAAATAGCTAACC 2017
Db 6706	AATGGCCCATGACGCCATGATAATTACTAATCTAGGAGAAGTTAATAAATAGCTAACC 6765
Qy 2018	AACATGATTACAAATTATTAGAGTCATCGTTTCAAAATGGTATCGGTTTTCGACATCCCA 2077

Db 6766 AACATGATTAAACAATTATTAGAGGTATCGTTCAAAATGGTATGCGTTTGGACACATCCA 6825
Qy 2078 CTATATATCCGTCGTCTGTCGCACTCTGATGCCAATCCAGAAATTCCTAGCGGATT 2137
Db 6826 CTATATATCCGTCGTCTGTCGCACTCTGATGCCAATCCAGAAATTCCTAGCGGATT 6885
Qy 2138 CCAGAAGTTTCTCAGAGTCGGAAGTTGACACAGACATTAACGAATCGGCACAGATGGTCAT 2197
Db 6886 CCAGAAGTTTCTCAGAGTCGGAAGTTGACACAGACATTAACGAATCGGCACAGATGGTCAT 6945
Qy 2198 AACCTGAAGGAAGATCTGATTCGTTAACTGCTTACGTTAAGACCGAAGCGCTCGTAT 2257
Db 6946 AACCTGAAGGAAGATCTGATTCGTTAACTGCTTACGTTAAGACCGAAGCGCTCGTAT 7005
Qy 2258 AACAGATCGGATGATGACAGACCAATCAACATGCGACCTGCCATTCGTACCTGTACAGTCA 2317
Db 7006 AACAGATCGGATGATGACAGACCAATCAACATGCGACCTGCCATTCGTACCTGTACAGTCA 7065
Qy 2318 AGGATGGTAGAAATGTTGTCGTCCTTGACACGAATATACGCAATTTGCGCTGCATATT 2377
Db 7066 AGGATGGTAGAAATGTTGTCGTCCTTGACACGAATATACGCAATTTGCGCTGCATATT 7125
Qy 2378 CAAACAGCTCTCTACGATAGGGACAAATCGCATCGTGAACGTTTGGGCTTCTACCG 2437
Db 7126 CAAACAGCTCTCTACGATAGGGACAAATCGCATCGTGAACGTTTGGGCTTCTACCG 7185
Qy 2438 ATTTAGCAGTTTGATACACTTTCTCTAAGTATCCACCTGAATCAATAACGCGCAAAATAG 2497
Db 7186 ATTTAGCAGTTTGATACACTTTCTCTAAGTATCCACCTGAATCAATAACGCGCAAAATAG 7245
Qy 2498 AGAAAAATGACCATGTGTAGCGGCCAATCTGATCCACCTGAGATGCATAATCTAGTA 2557
Db 7246 AGAAAAATGACCATGTGTAGCGGCCAATCTGATCCACCTGAGATGCATAATCTAGTA 7305
Qy 2558 GAATCTCTCGGTATCAAAATTCACCTCCACCTTCCACTCACCGGTGTCCATTCATGGC 2617
Db 7306 GAATCTCTCGGTATCAAAATTCACCTCCACCTTCCACTCACCGGTGTCCATTCATGGC 7365
Qy 2618 TGAATCTGCTCTGTCGTGTGACATGACACATCATCTCAATTCGAAATAGGGCCCAT 2677
Db 7366 TGAATCTGCTCTGTCGTGTGACATGACACATCATCTCAATTCGAAATAGGGCCCAT 7425
Qy 2678 CAGTCTGACGACCAAGAGGCCATAACACCAATAGCCTTAACATCATCCCATATTAT 2737
Db 7426 CAGTCTGACGACCAAGAGGCCATAACACCAATAGCCTTAACATCATCCCATATTAT 7485
Qy 2738 CCAATATTGCTTCCCTTAATTCATGAACAATCTTCATTTCTCTCTAGTCATATTATTA 2797
Db 7486 CCAATATTGCTTCCCTTAATTCATGAACAATCTTCATTTCTCTCTAGTCATATTATTA 7545
Qy 2798 TTGGTCCATTCATCTCTCCCTTTTCAGATAATTTAGATTTGCTTTTCTTAATA 2857
Db 7546 TTGGTCCATTCATCTCTCCCTTTTCAGATAATTTAGATTTGCTTTTCTTAATA 7605
Qy 2858 AGAATATTGGAGACCGCTTCTTATTACGCTATTAACTCTGCTTCTCTAAGCATCC 2917
Db 7606 AGAATATTGGAGACCGCTTCTTATTACGCTATTAACTCTGCTTCTCTAAGCATCC 7665
Qy 2918 TTCAATCTTTTAATAAAGATTAAGATCTAATCTTCAACAACTGGCCGTTGTGA 2977
Db 7666 TTCAATCTTTTAATAAAGATTAAGATCTAATCTTCAACAACTGGCCGTTGTGTGA 7725
Qy 2978 ACTACTCTTTAATAAATTTTCCGTTCCCAATTCACATTCGCAATATAGAAATC 3037
Db 7726 ACTACTCTTTAATAAATTTTCCGTTCCCAATTCACATTCGCAATATAGAAATC 7785
Qy 3038 CATCTTCATCGGCTTTTGTGATCATCTGTATGAATCAATCGCTTCTTCTGTGTCAT 3097
Db 7786 CATCTTCATCGGCTTTTGTGATCATCTGTATGAATCAATCGCTTCTTCTGTGTCAT 7845
Qy 3098 CAAAGGTTTAATTTTATGATTTCTTTTAAACAAACCAACCATAGGAGATTAACCTTTTAC 3157
Db 7846 CAAAGGTTTAATTTTATGATTTCTTTTAAACAAACCAACCATAGGAGATTAACCTTTTAC 7905

Qy 3158 GGTGTAACCTTCCCTCCAAATCAGACAAACGTTTCAAAATTCCTTCTCATCATCGGTCA 3217
Db 7906 GGTGTAACCTTCCCTCCAAATCAGACAAACGTTTCAAAATTCCTTCTCATCATCGGTCA 7965
Qy 3218 TAAATCCGTATCCCTTTCAGGATATTTTTCAGTTTTCGTCATTTGTCGCAATTCGCAATTCG 3277
Db 7966 TAAATCCGTATCCCTTTCAGGATATTTTTCAGTTTTCGTCATTTGTCGCAATTCGCAATTCG 8025
Qy 3278 ATTATATTTTATTTTTCGATTTTTCGTTTATTTTAAACGTCCTCA 3318
Db 8026 ATTATATTTTATTTTTCGATTTTTCGTTTATTTTAAACGTCCTCA 8066

RESULT 15

5304637-21/c
; Patent No. 5304637
; APPLICANT: DORSSERS, LAMBERTUS C.J.; WAGEMAKER, GERARD; VOS,
; VONNE J.; VAN LEEN, ROBERT W.; PERSOON, MARIA L.N.
; TITLE OF INVENTION: EXPRESSION AND PURIFICATION OF HUMAN
; INTERLEUKIN-3 AND MUTINS THEREOF
; NUMBER OF SEQUENCES: 30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/494,182
; FILING DATE: 13-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 249,184
; FILING DATE: 16-AUG-1988
; SEQ ID NO: 21:
; LENGTH: 7326
5304637-21

Query Match 29.0%; Score 1491.4; DB 6; Length 7326;

Best Local Similarity 98.0%; Pred. No. 0;
Matches 1510; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
Qy 1778 TTCTTATCTTGATAATAAGGGTAACTATTGCGGATAGACTGTAACTTCTCAGGCATA 1837
Db 6367 TACTCTTTTATATATCCCGACTGGCAATGCGGATAGACTGTAACTTCTCAGGCATA 6308
Qy 1838 AAATCCCTTTCATTTCTTAATGTAATCTATTAACCTTATTAATTAATTAATTCGCTCAT 1897
Db 6307 AAATCCCTTTCATTTCTTAATGTAATCTATTAACCTTATTAATTAATTCGCTCAT 6248
Qy 1898 AATTAATCTCTTTCTTATTAGCAAAATGGCCGATTTAAGCACACCTTTTATTCGCTT 1957
Db 6247 AATTAATCTCTTTCTTATTAGCAAAATGGCCGATTTAAGCACACCTTTTATTCGCTT 6188
Qy 1958 AATGCGCCATGACAGCCATGATAATTAATACTAGGAGAAAGTTAATAATACGTAACC 2017
Db 6187 AATGCGCCATGACAGCCATGATAATTAATACTAGGAGAAAGTTAATAATACGTAACC 6128
Qy 2018 AACATGATTAACAATTTATAGAGTCATCGTTCAAAATGGTATGCTTTTGACACATCCA 2077
Db 6127 AACATGATTAACAATTTATAGAGTCATCGTTCAAAATGGTATGCTTTTGACACATCCA 6068
Qy 2078 CTATATATCCGTCGTCTGCTCTCCACTCTCTGAATCCCAATTCAGAAATTCCTAGCGATT 2137
Db 6067 CTATATATCCGTCGTCTGCTCTCCACTCTCTGAATCCCAATTCAGAAATTCCTAGCGATT 6008
Qy 2138 CAGAAGTTTCTCAGAGTCGGAAGTTGACAGACATTAACGAGCTGGCAGACATGGTCAT 2197
Db 6007 CAGAAGTTTCTCAGAGTCGGAAGTTGACAGACATTAACGAGCTGGCAGACATGGTCAT 5948
Qy 2198 AACCTGAAGGAAGATCTGATTCGTTAACTGCTTAAAGTAAAGACCGAAGCGCTCGTCGAT 2257
Db 5947 AACCTGAAGGAAGATCTGATTCGTTAACTGCTTAAAGTAAAGACCGAAGCGCTCGTCGAT 5888
Qy 2258 AACAGATGCGATGATGACAGACCAATCAACATGGCACTTGGCAATTCCTACCTGTACAGTCA 2317
Db 5887 AACAGATGCGATGATGACAGACCAATCAACATGGCACTTGGCAATTCCTACCTGTACAGTCA 5828
Qy 2318 AGGATGGTAAAGATTTGTCGCTTCTTGACACGAATATTAACGCAATTCCTGTCATATT 2377

5827	AGGATGGTAAAGATGTTGTGGGTCCTTGGACAGAAATATTAAGCCATTTGGCTGCTATATT	5769
2378	CAAAACAGCTCTTCTACGATAAAGGCAAAATCGCATCGTGGAAACGTTTGGGCTTCTACCG	2437
5767	CAAAACAGCTCTTCTACGATAAAGGCAAAATCGCATCGTGGAAACGTTTGGGCTTCTACCG	5708
2438	ATTTAGCAGTTTGATACACCTTCTTAAGTATCCACCTGAATCATAAATCGCGAAANTAG	2497
5707	ATTTAGCAGTTTGATACACCTTCTTAAGTATCCACCTGAATCATAAATCGCGAAANTAG	5648
2498	AGAAAAATTCACCATGTGTAAGCGGCAATCTCATTTCCACCTGAGATGCATAATCTAGTA	2557
5647	AGAAAAATTCACCATGTGTAATCGGCCAATCTGATTTCCACCTGAGATGCATAATCTAGTA	5588
2558	GAATCTCTTCGCTATCAAAAATTCACCTTCACCTTCACCTGCGGTTGTCATTCATGCGC	2617
5587	GAATCTCTTCGCTATCAAAAATTCACCTTCACCTTCACCTGCGGTTGTCATTCATGCGC	5528
2618	TGAATCTGCTCTCTCTGTTGACATGACACACATCATCTCAATATCCGAATAGGGGCCAT	2677
5527	TGAATCTGCTCTCTCTGTTGACATGACACACATCATCTCAATATCCGAATAGGGGCCAT	5468
2678	CAGTCTGACCAACCAAGAGAGCCATAACACCAATAGCCTTAACATCATCCCCATATTTAT	2737
5467	CAGTCTGACCAACCAAGAGAGCCATAACACCAATAGCCTTAACATCATCCCCATATTTAT	5408
2738	CCAATATTCGTTCCCTTAATTTTCATGAACAATCTTCATTTCTCTCTAGTCATTATTA	2797
5407	CCAATATTCGTTCCCTTAATTTTCATGAACAATCTTCATTTCTCTCTAGTCATTATTA	5348
2798	TTGGTCCAAATTCACATTTCTCATTCCTCTTTTCAGATAATTTAGATTTGCTTTTCTAAATA	2857
5347	TTGGTCCAAATTCACATTTCTCATTCCTCTTTTCAGATAATTTAGATTTGCTTTTCTAAATA	5288
2858	AGAAATTTGGAGAGCACCGTTCTTATTCAGCTATTTAATCACTGCTCTTCTTAAGCATCC	2917
5287	AGAAATTTGGAGAGCACCGTTCTTATTCAGCTATTTAATAACTCGTCTTCTTAAGCATCC	5228
2918	TTCAATCCTTTTAATAACAATTAAGATCTTAATCTTCAACAACTGGCCGTTTGTGTA	2977
5227	TTCAATCCTTTTAATAACAATTAAGATCTTAATCTTCAACAACTGGCCGTTTGTGTA	5168
2978	ACTACTCTTTAATAAATAATTTTCCGTTCCCAATTCCAATTCGAATAATAGAAAAATC	3037
5167	ACTACTCTTTAATAAATAATTTTCCGTTCCCAATTCCAATTCGAATAATAGAAAAATC	5108
3038	CATCTTCATCGCTTTTTCGTCATCACTGTATCAATCAAAATCGCTTCTTCTGTGTTCAT	3097
5107	CATCTTCATCGCTTTTTCGTCATCACTGTATCAAAATCGCTTCTTCTGTGTTCAT	5048
3098	CAAGGTTTAAATTTTATGTATTTCTTTTAAACCAACCAATAGGAGATTAACCTTTTAC	3157
5047	CAAGGTTTAAATTTTATGTATTTCTTTTAAACCAACCAATAGGAGATTAACCTTTTAC	4988
3158	GGGTAAACCTTCTCCAAAATCAGACAAACGTTTCAAAATTCCTTTCTTCATCGGTCA	3217
4987	GGGTAAACCTTCTCCAAAATCAGACAAACGTTTCAAAATTCCTTTCTTCATCGGTCA	4928
3218	TAAATTCGGTATCCCTTTTACAGATATTTTGCAGTTTCGTCAAATGCGGATGTATATCCG	3277
4927	TAAATTCGGTATCCCTTTTACAGATATTTTGCAGTTTCGTCAAATGCGGATGTATATCCG	4868
3278	ATTTATATTTATTTTTCGGTATTTTTTATTAATAACGTCCTCA	3318
4867	ATTTATATTTATTTTTCGGTCGAATCATTTGAATTTTACA	4827

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM nucleic - nucleic search, using sw model

Run on: September 23, 2004, 02:36:16 ; Search time 1169.56 Seconds
(without alignments)
18677.278 Million cell updates/sec

Title: US-10-030-390-1

Perfect score: 5142

Sequence: 1 gaattcgagcgcgtaccgcg.....ccgtctgaagtcctcttaag 5142

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 29Jan04: *
1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002s: *
7: Geneseqn2003as: *
8: Geneseqn2003bs: *
9: Geneseqn2003cs: *
10: Geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5142	100.0	5142	5	Aaf29645 Plasmid p
2	4232.8	82.3	10929	6	Abx98591 pep25 vec
3	4232.8	82.3	10929	8	Acid13842 Plasmid p
C 4	2467.2	48.0	8119	2	Aat33477 Subtilisi
C 5	1650.4	32.1	2059	1	Aan70271 DNA fragm
6	1636.8	31.8	3792	2	Aa48463 Plasmid p
7	1635.2	31.8	5234	2	Aa48464 Plasmid p
8	1635.2	31.8	6722	2	Aa48465 Plasmid p
C 9	1493	29.0	7336	1	Aan91856 Plasmid p
10	1493	29.0	8808	2	Aat32227 Plasmid p
11	1493	29.0	9144	2	Aat18951 Plasmid p
12	1493	29.0	10140	2	Aat32232 Plasmid p
13	1490.6	29.0	6661	7	Aad49392 pMOL1995 p
14	1487.6	28.9	5302	3	Aaa98012 Expressio
15	1487.6	28.9	5767	3	Aaa98014 Expressio
16	1481	28.8	2763	2	Aa48044 Lactobaci
C 17	1479	28.8	5609	6	Aad23867 pCFC717 D
C 18	1479	28.8	5609	6	Aad23868 pEB200 DN
C 19	1479	28.8	7026	6	Aad23871 pEB303 pl
C 20	1479	28.8	7344	6	Aad23869 pEB301 ch
21	1476.2	28.7	2763	2	Aa48037 Plasmid p
C 22	1464	28.5	7534	1	Aan80664 Sequence
C 23	1363	26.5	7456	2	Aaql0686 Plasmid p

ALIGNMENTS

RESULT 1

Aaf29645

ID Aaf29645 standard; DNA; 5142 BP.

XX AC Aaf29645;

DT 10-APR-2001 (first entry)

DE Plasmid pL2mTFF1v1.

XX Mouse TFF1; mTFF; trefoil protein; antiinflammatory; gastrointestinal;
XX anti-ulcer; peptide therapy; gastrointestinal disease; acute colitis;
XX Crohn's disease; ulcerative colitis; plasmid pL2mTFF1v1;
XX recombinant vector; ds.
XX Synthetic.
OS
XX WO200102570-A1.
XX
XX 11-JAN-2001.
XX
XX 05-JUL-2000; 2000WO-EP006343.
XX
XX 05-JUL-1999; 99EP-00870143.
XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
XX Hans WC, Steidler L, Remaut ER;
XX
XX WPI; 2001-138142/14.
XX
XX Recombinant Lactococcus lactis for delivering a trefoil peptide useful
XX for treating acute or chronic gastrointestinal inflammatory diseases or
XX disorders, e.g. acute or ulcerative colitis, acute flare-ups of Crohn's
XX disease.
XX
XX Claim 18, Fig 1b; 4app; English.
XX
XX The present sequence is a recombinant vector which may be used to
XX generate a recombinant Lactococcus lactis capable of delivering a trefoil
XX peptide in vivo. The recombinant microorganism is useful for
XX manufacturing an agent for the delivery of a trefoil peptide to the
XX gastrointestinal tract, and for treating gastric or intestinal diseases
XX or disorders, or lesions caused by gastric or intestinal diseases or
XX disorders. The microorganism may also be used for preparing medicament to
XX be used for treating gastric and /or gastrointestinal diseases or

disorders, acute gastrointestinal inflammatory diseases (e.g., acute colitis, acute flare-ups of Crohn's diseases, or ulcerative colitis), and chronic and spontaneously recurring diseases of the gastrointestinal tract comprising Crohn's disease (enteritis regionalis) and ulcerative colitis (colitis ulcerosa). Disease states which can be treated by the method or compositions comprising the recombinant microorganism or trefoil peptides include disorders of and damage to the alimentary canal, including the mouth, oesophagus, stomach and large and small intestine, as well as for the protection and treatment of tissues that lie outside the alimentary canal.

XX SQ Sequence 5142 BP; 1617 A; 954 C; 927 G; 1644 T; 0 U; 0 Other;

Query Match 100.0%; Score 5142; DB 5; Length 5142;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GAATTCGAGCTCGGTACCGGGGATCTCGATCCCGGAATTAATACGACTCAGCTATAGG	60
DB	1	GAATTCGAGCTCGGTACCGGGGATCTCGATCCCGGAATTAATACGACTCAGCTATAGG	60
QY	61	GAGACCAACAGCGTTCCCTCTAGAAATAATTTGTTTAACTTTAAGAAGGAGATATACA	120
DB	61	GAGACCAACAGCGTTCCCTCTAGAAATAATTTGTTTAACTTTAAGAAGGAGATATACA	120
QY	121	TATGAAAAAAGATTATCTCAGCTATTTTAAATGCTACAGTATATCTTCTGCTGAGC	180
DB	121	TATGAAAAAAGATTATCTCAGCTATTTTAAATGCTACAGTATATCTTCTGCTGAGC	180
QY	181	CCCGTTGTCAGGTGTTTACGCCCGCCAGGCCCGCCAGGCCCGCCAGGAAACATGTATCAT	240
DB	181	CCCGTTGTCAGGTGTTTACGCCCGCCAGGCCCGCCAGGCCCGCCAGGAAACATGTATCAT	240
QY	241	GGCCCCCGGAGAGGATTAATTTGTTGCTCCCGGTGTACCGCCCGAGCAGTGCACGGA	300
DB	241	GGCCCCCGGAGAGGATTAATTTGTTGCTCCCGGTGTACCGCCCGAGCAGTGCACGGA	300
QY	301	GAGAGGTTGCTGTTTGTATGACAGTGTCCGGGATTCCTGGTGTCTCCACCCCATGGC	360
DB	301	GAGAGGTTGCTGTTTGTATGACAGTGTCCGGGATTCCTGGTGTCTCCACCCCATGGC	360
QY	361	CATGAGAACACTCAAGAAGAAATGTCCCTTCTAACTAGTAGATCCGCTGCTAACAA	420
DB	361	CATGAGAACACTCAAGAAGAAATGTCCCTTCTAACTAGTAGATCCGCTGCTAACAA	420
QY	421	AGCCGGAAGGAGCTGAGTTGGCTGCTGCCCGCTGAGCAATTAAGTAAACCCCT	480
DB	421	AGCCGGAAGGAGCTGAGTTGGCTGCTGCCCGCTGAGCAATTAAGTAAACCCCT	480
QY	481	TGGGGCCCTCTAAACCGGTCTTTGAGGGTCTTTTGGCTGAAAGGAGGAATATATCCGGATG	540
DB	481	TGGGGCCCTCTAAACCGGTCTTTGAGGGTCTTTTGGCTGAAAGGAGGAATATATCCGGATG	540
QY	541	ACCTGCGAGGATGCAAGCTGGGCTGCTGCCCGCTGAGCAATTAAGTAAACCCCT	600
DB	541	ACCTGCGAGGATGCAAGCTGGGCTGCTGCCCGCTGAGCAATTAAGTAAACCCCT	600
QY	601	CCTGGCGTTACCCAACTTAATTCGGCTTGCAGCACATCCCGCTTTCGCCAGCTGATTTAC	660
DB	601	CCTGGCGTTACCCAACTTAATTCGGCTTGCAGCACATCCCGCTTTCGCCAGCTGATTTAC	660
QY	661	TTTTTGATTTTACAACTGATATCTATGTAATTCGCTCTTTTAAAGTGGGACAA	720
DB	661	TTTTTGATTTTACAACTGATATCTATGTAATTCGCTCTTTTAAAGTGGGACAA	720
QY	721	ATGTAGGCAATTTTCGCTCTTTTCGGCGAGGCTAGTTTACCTTAAGTATGATGACT	780
DB	721	ATGTAGGCAATTTTCGCTCTTTTCGGCGAGGCTAGTTTACCTTAAGTATGATGACT	780
QY	781	GGTTTTAAGCGCAAAAAGTTGCTTTTTCGTAACCTATTAATGATCGTTTTAGAAAACC	840
DB	781	GGTTTTAAGCGCAAAAAGTTGCTTTTTCGTAACCTATTAATGATCGTTTTAGAAAACC	840

QY	841	GACTGTAAAAAGTACAGTCGGCATTATCTCATATTATAAAAGCCAGTCATTAGGCTATC	900
DB	841	GACTGTAAAAAGTACAGTCGGCATTATCTCATATTATAAAAGCCAGTCATTAGGCTATC	900
QY	901	TGCAATATTCCTGAATAGAGTTCATAAACAATCCTGCATGATAACCATCACAACAGATG	960
DB	901	TGCAATATTCCTGAATAGAGTTCATAAACAATCCTGCATGATAACCATCACAACAGATG	960
QY	961	ATGTACTCTGTAAGATAGCGGTAAATATATCAATTTACCTTTATTAAATGAATTTTCCTGC	1020
DB	961	ATGTACTCTGTAAGATAGCGGTAAATATATCAATTTACCTTTATTAAATGAATTTTCCTGC	1020
QY	1021	TGTAATATGGGTAGAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1080
DB	1021	TGTAATATGGGTAGAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1080
QY	1081	TGAAGTCCCATGGAATAATAGAAAGAGAAAGCAATTTTCAGGTATAGGTGTTTTGGGAAA	1140
DB	1081	TGAAGTCCCATGGAATAATAGAAAGAGAAAGCAATTTTCAGGTATAGGTGTTTTGGGAAA	1140
QY	1141	CAATTTCCCGCAACCATTAATTTCTACATCAGAAAGGTATTAATCATATAAACTCTTT	1200
DB	1141	CAATTTCCCGCAACCATTAATTTCTACATCAGAAAGGTATTAATCATATAAACTCTTT	1200
QY	1201	GAGTCATTTCTACAGGATCCCAATACAGAGAGTGTGTAGTATACCATCAAAAT	1260
DB	1201	GAGTCATTTCTACAGGATCCCAATACAGAGAGTGTGTAGTATACCATCAAAAT	1260
QY	1261	TGTAATAAGTGGCTCTAACTTATCCCAATAACCTTCTCCGTCGCTATTTGTAACAGT	1320
DB	1261	TGTAATAAGTGGCTCTAACTTATCCCAATAACCTTCTCCGTCGCTATTTGTAACAGT	1320
QY	1321	TCTAAAAGCTGTATTTGATTTATCACCCCTTGTGCACTAAAGAAATAAATCAGGGTAAAA	1380
DB	1321	TCTAAAAGCTGTATTTGATTTATCACCCCTTGTGCACTAAAGAAATAAATCAGGGTAAAA	1380
QY	1381	TTTATATCTTCTCTGTTTATGTTTCGGTATAAACAACACTAATCAATTTCTGTTGTTAT	1440
DB	1381	TTTATATCTTCTCTGTTTATGTTTCGGTATAAACAACACTAATCAATTTCTGTTGTTAT	1440
QY	1441	ACTAAAAGCTGTATTTGATTTCAATAATTAATTAATTAATTTCTCTTTTCTTTCCAAATGTC	1500
DB	1441	ACTAAAAGCTGTATTTGATTTCAATAATTAATTAATTAATTTCTCTTTTCTTTCCAAATGTC	1500
QY	1501	TAAATCAATTTTATTAAGTTCATTTGATTCGCTCTTAATTTTATCTTAAGTGAATTT	1560
DB	1501	TAAATCAATTTTATTAAGTTCATTTGATTCGCTCTTAATTTTATCTTAAGTGAATTT	1560
QY	1561	TAGGAGGCTTACTTGTCTGCTTCTTCTTCAATTAAGTCAATCCTTTTAAAAAGTCAATTT	1620
DB	1561	TAGGAGGCTTACTTGTCTGCTTCTTCTTCAATTAAGTCAATCCTTTTAAAAAGTCAATTT	1620
QY	1621	ACTGTAAACATAAATATATTTTAAAAATATCCCACTTATCCAAATTTGTTGTTGAA	1680
DB	1621	ACTGTAAACATAAATATATTTTAAAAATATCCCACTTATCCCAATTTGTTGTTGAA	1680
QY	1681	CTAATGGGTGCTTTAGTTGAAGATTAAGACCAATTAAGACCAATTAAGACCAATTAAG	1740
DB	1681	CTAATGGGTGCTTTAGTTGAAGATTAAGACCAATTAAGACCAATTAAGACCAATTAAG	1740
QY	1741	TTTTTAAGGATTTGAGCGTAGCGAAAAATCCTTTTCTTCTTATCTTGAATAAAGGGTA	1800
DB	1741	TTTTTAAGGATTTGAGCGTAGCGAAAAATCCTTTTCTTCTTATCTTGAATAAAGGGTA	1800
QY	1801	ACTATTCGCGGATAGACTGTAACTTCTCAGCATAAAAATCCCTTTTCAATTTCTAATG	1860
DB	1801	ACTATTCGCGGATAGACTGTAACTTCTCAGCATAAAAATCCCTTTTCAATTTCTAATG	1860
QY	1861	TAAATCTATTAACCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1920
DB	1861	TAAATCTATTAACCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1920
QY	1921	CAAAATGCGCGATTTTAAAGCACACCCCTTTATTCGTTTAAATGCGCCATGACAGCATGATA	1980

Db	1921	CAAAATGGCCGATTTAAAGCACACCCCTTTATTCCGTTAATGGCCATGACGCCATGATA	1980
Qy	1981	ATTACTAATACTAGGAGAGTTAATAAATACGTAAACCAACATGATTAAACAAATTATTAGAG	2040
Db	1981	ATTACTAATACTAGGAGAGTTAATAAATACGTAAACCAACATGATTAAACAAATTATTAGAG	2040
Qy	2041	GTCAATCGTTCAAAATGGTATGCGTTTGTGACACATCCACTATATATCCGTGTCGTTCTGTC	2100
Db	2041	GTCAATCGTTCAAAATGGTATGCGTTTGTGACACATCCACTATATATCCGTGTCGTTCTGTC	2100
Qy	2101	CATCTCTGAATCCGATTCAGAAATTTCTCTAGCGATTCCAGAGTTTCTCAGAGTCGGA	2160
Db	2101	CATCTCTGAATCCGATTCAGAAATTTCTCTAGCGATTCCAGAGTTTCTCAGAGTCGGA	2160
Qy	2161	AGTTGACCAACATTAACGAATCGCACAGATGTCATAACCTGGAAGAAAGATCTGATTC	2220
Db	2161	AGTTGACCAACATTAACGAATCGCACAGATGTCATAACCTGGAAGAAAGATCTGATTC	2220
Qy	2221	TTAACTGCTTCAGTTAAGACCGAAGCGCTCGTCGTATTAACAGATGCGATGCAGACCA	2280
Db	2221	TTAACTGCTTCAGTTAAGACCGAAGCGCTCGTCGTATTAACAGATGCGATGCAGACCA	2280
Qy	2281	ATCAACATGCGACCTGTCATTTGCTGCTGCAATTTCAACAGTCAAGATGCTGAAATGTTTCGCT	2340
Db	2281	ATCAACATGCGACCTGTCATTTGCTGCTGCAATTTCAACAGTCAAGATGCTGAAATGTTTCGCT	2340
Qy	2341	CGTTGCAACGAATATTAACGCAATTTGCTGCAATTTCAACAGTCTTCTACGATTAAGS	2400
Db	2341	CGTTGCAACGAATATTAACGCAATTTGCTGCTGCAATTTCAACAGTCTTCTACGATTAAGS	2400
Qy	2401	GCACAAATCGCATCGTGGAAAGTTGGGCTTCTACCGATTTAGCAGTTTGATACATTTTC	2460
Db	2401	GCACAAATCGCATCGTGGAAAGTTGGGCTTCTACCGATTTAGCAGTTTGATACATTTTC	2460
Qy	2461	TCTAAGTATCCACTGTAATCATAAATCGGCAAAATAGAGAAAATTGACCAATGTGAAGC	2520
Db	2461	TCTAAGTATCCACTGTAATCATAAATCGGCAAAATAGAGAAAATTGACCAATGTGAAGC	2520
Qy	2521	GGCCAAATCTGATTTCCACTGAGATGCATATCTAGTAGAATCTCTTCGCTATCABAATTC	2580
Db	2521	GGCCAAATCTGATTTCCACTGAGATGCATATCTAGTAGAATCTCTTCGCTATCABAATTC	2580
Qy	2581	ACTTCCACCTTCCACTCACCGGTTGTCATTCATGCGTGAACCTGCTTCTCTCTGTTGAC	2640
Db	2581	ACTTCCACCTTCCACTCACCGGTTGTCATTCATGCGTGAACCTGCTTCTCTCTGTTGAC	2640
Qy	2641	ATGACACATCATCTCATAATCCGATAGGGCCCACTAGTCTGACACCAAGAGAGCCA	2700
Db	2641	ATGACACATCATCTCATAATCCGATAGGGCCCACTAGTCTGACACCAAGAGAGCCA	2700
Qy	2701	TAAACACCAATAGCCTTAACATCATCCCATATTTATCCCAATTCGTTCTTAAATTTCA	2760
Db	2701	TAAACACCAATAGCCTTAACATCATCCCATATTTATCCCAATTCGTTCTTAAATTTCA	2760
Qy	2761	TGAACAATCTTCAATCTTCTCTAGTCAATTAATTTGGTCCATTCACATTTCTCATTT	2820
Db	2761	TGAACAATCTTCAATCTTCTCTAGTCAATTAATTTGGTCCATTCACATTTCTCATTT	2820
Qy	2821	CCCTTTTTCAGATATTTAGATTTGCTTTTCTTAATAGATAATTTGGAGAGACCGGTTTC	2880
Db	2821	CCCTTTTTCAGATATTTAGATTTGCTTTTCTTAATAGATAATTTGGAGAGACCGGTTTC	2880
Qy	2881	TTAATTCAGCTATTAATAACTCGTCTTCTTAAGCATCTTCAATCTTTTAAATTAACAATTA	2940
Db	2881	TTAATTCAGCTATTAATAACTCGTCTTCTTAAGCATCTTCAATCTTTTAAATTAACAATTA	2940
Qy	2941	TAGCATCTAATCTTCAACAACTGGCCGGTTTGTGAACTACTCTTTAAATAAATAATTT	3000
Db	2941	TAGCATCTAATCTTCAACAACTGGCCGGTTTGTGAACTACTCTTTAAATAAATAATTT	3000
Qy	3001	TTCCGTTCCCAATTCACATTCGAATAATAGAAAATCCATCTTTCATCGGCTTTTTCGTCA	3060

QY 4141 ATTTTATATATCTGACTCAATTTCTTAAGTATGGAACAAAAATTAGAGCTTTGGG 4200
Db 4141 ATTTTATATATCTGACTCAATTTCTTAAGTATGGAACAAAAATTAGAGCTTTGGG 4200
QY 4201 COTATCTATGCTGTGCTGCTTCTTACAGCATATGACGAAAAAAGATAAAGATACATG 4260
Db 4201 COTATCTATGCTGTGCTGCTTCTTACAGCATATGACGAAAAAAGATAAAGATACATG 4260
QY 4261 GAATAGTAGTATGATATACGAATAGGAACACATATATAAACCACACTATCAGTTAT 4320
Db 4261 GAATAGTAGTATGATATACGAATAGGAACACATATATAAACCACACTATCAGTTAT 4320
QY 4321 ATATATTCAGCAAAATCCTGTAAACAATAAGAAAGCTTAGGAACAAGATTAAGCGAAATTT 4380
Db 4321 ATATATTCAGCAAAATCCTGTAAACAATAAGAAAGCTTAGGAACAAGATTAAGCGAAATTT 4380
QY 4381 GGGGAATAGTTCAGTTGCTCATGTTGAGATACCTGATATATCAAAAGGTTTCATATGAATA 4440
Db 4381 GGGGAATAGTTCAGTTGCTCATGTTGAGATACCTGATATATCAAAAGGTTTCATATGAATA 4440
QY 4441 TTTGACTCATGAATCAAAAGGAGCTTATTGCTAAGATAAACAATATATACGACAAAAAGA 4500
Db 4441 TTTGACTCATGAATCAAAAGGAGCTTATTGCTAAGATAAACAATATATACGACAAAAAGA 4500
QY 4501 TATTTTGAACATTAATGATTTTGTATGATGACCGCTTATATAACCTTGATGAAAGCCAAAA 4560
Db 4501 TATTTTGAACATTAATGATTTTGTATGATGACCGCTTATATAACCTTGATGAAAGCCAAAA 4560
QY 4561 AAGCAATAGGAATTTACCTTTAGATATAGTGGTACCTATATTTGGTAAATACAAA 4620
Db 4561 AAGCAATAGGAATTTACCTTTAGATATAGTGGTACCTATATTTGGTAAATACAAA 4620
QY 4621 AGATTTAATGCTTTTATTCGCTTTAGGGAGCGAGTTTGGAAATTTAAATACGAATGA 4680
Db 4621 AGATTTAATGCTTTTATTCGCTTTAGGGAGCGAGTTTGGAAATTTAAATACGAATGA 4680
QY 4681 TGTAAAGATATGTTTCAACAACCTTAGCGCTTAGATATATGTTTGGGCAATTA 4740
Db 4681 TGTAAAGATATGTTTCAACAACCTTAGCGCTTTAGATATATGTTTGGGCAATTA 4740
QY 4741 TCAGTGTGATATAGCAAGTATGCAAGGTTCTTGATCTGAAACGGGGGAAATAAA 4800
Db 4741 TCAGTGTGATATAGCAAGTATGCAAGGTTCTTGATCTGAAACGGGGGAAATAAA 4800
QY 4801 ATGACAAACAAGAAAAGAGTATTTGCTGAAATAGGGAATTAATAAAGAAATTAAG 4860
Db 4801 ATGACAAACAAGAAAAGAGTATTTGCTGAAATAGGGAATTAATAAAGAAATTAAG 4860
QY 4861 GACTTAAAGAGCGTATTGAAGATACAGAGAAATGGAAGTTGAATTAAGTACAAACATA 4920
Db 4861 GACTTAAAGAGCGTATTGAAGATACAGAGAAATGGAAGTTGAATTAAGTACAAACATA 4920
QY 4921 GATTTATTGAGGAGGAGGATTTGAATTAATAAAGCCCTTCGACGAAGTCGCGACT 4980
Db 4921 GATTTATTGAGGAGGAGGATTTGAATTAATAAAGCCCTTCGACGAAGTCGCGACT 4980
QY 4981 TCGTTCTTTTTCCTCTCGTATGATAGTTTCAAAATCGTTCTTTTATAGTTCTAA 5040
Db 4981 TCGTTCTTTTTCCTCTCGTATGATAGTTTCAAAATCGTTCTTTTATAGTTCTAA 5040
QY 5041 ATCGTGTTTTTCCTGGAATTTGCTGTTTATTCCTTTTACCTTGTCTACAAACCCCTTAAA 5100
Db 5041 ATCGTGTTTTTCCTGGAATTTGCTGTTTATTCCTTTTACCTTGTCTACAAACCCCTTAAA 5100
QY 5101 AACGTTTTTAAAGGCTTTTAAAGCGGCTGTAGCTTCCTTAG 5142
Db 5101 AACGTTTTTAAAGGCTTTTAAAGCGGCTGTAGCTTCCTTAG 5142

RESULT 2
ABK98591
ID ABK98591 standard; DNA; 10929 BP.

XX ABK98591;
XX AC
XX 07-AUG-2003 (revised)
DT 21-OCT-2002 (first entry)
XX
DE pBP25 vector fragment containing XylR/XyIO/CP25 sequences.
XX
KW ds; promoter; gram positive bacteria; fusion promoter; TS; CP25; P32;
KW P59; P1P2; PL; xyIO; tetO; trpO; malO; lambdaclO; cellular proliferation;
KW antibiotic; vector.
XX
OS Eubacteria.
OS Bacteriophage lambda.
OS Escherichia coli.
OS Synthetic.
XX
PN WO200251982-A2.
XX
XX 04-JUL-2002.
XX
XX 21-DEC-2001; 2001WO-US050250.
XX
XX 27-DEC-2000; 2000US-0259434P.
PR 06-SEP-2001; 2001US-00948993.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Wall D, Gross M;
XX WPI; 2002-575374/61.
XX
PT Isolated nucleic acid comprises bacterial promoters modified to have
PT altered activity in at least one gram-positive organism, e.g. Bacillus
PT anthracis or Clostridium botulinum, useful for regulating gene expression
PT in bacteria.
XX
PS Example 1; Page 202-205; 246pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising a fusion
CC promoter comprising at least one promoter that is modified to have
CC altered activity in at least one gram-positive organism, or comprising
CC TS, CP25, P32, P59, P1P2 or PL linked to at least one operator consisting
CC of xyIO, tetO, trpO, malO or lambdaclO, where at least one operator is
CC positioned so binding of a repressor to an operator represses
CC transcription from the fusion promoter. Also included are vectors and
CC host cells comprising the fusion promoters, a method of identifying genes
CC involved in cellular proliferation or required for proliferation of a
CC prokaryotic cell using the vector, a method of identifying compounds that
CC inhibit the proliferation of a prokaryotic cell using the vector, a
CC method of identifying a compound that reduces the activity or level of a
CC gene product required for proliferation of a cell using the vector, a
CC compound identified by the methods, a method of inhibiting the activity
CC or expression of a gene in an operon required for proliferation using the
CC vector, manufacturing an antibiotic comprising using the vector or cell
CC and identifying a nucleic acid with promoter activity in Enterococcus
CC faecalis. The fusion promoters are useful for regulating nucleic acid or
CC polypeptide expression, particularly for regulating gene expression in
CC bacteria and for identifying proliferation-required genes or molecules
CC with potential antibiotic activity. The modified promoters are also
CC useful for replacing endogenous promoters to create cells with specific
CC regulatable genes. The present sequence is vector (or fragment)
CC incorporating a fusion promoter sequence of the invention. (Updated on 07
CC -AUG-2003 to correct OS field.)
XX
SQ Sequence 10929 BP; 3259 A; 2175 C; 2157 G; 3338 T; 0 U; 0 Other;

Query Match 82.3%; Score 4232.8; DB 6; Length 10929;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 4323; Conservative 0; Mismatches 17; Indels 10; Gaps 7;
QY 651 CTGATTTCACCTTTTGGCTTCTTACAACTGCATATGTAATCGTCCCTTTTA 710
|||||

Db 6491 CTGATTTCACTTTTGGTCAATCTACAAAGTGCATTAACCTCATATGTAAATCGCTCCTTTTAA 6550
 Qy 711 GGTGGCAAAATGTAGGCAATTTTCGCTCTTTCCGGGAGGCTAGTTACCTTTAAGTTAT 770
 Db 6551 GGTGGCAAAATGTAGGCAATTTTCGCTCTTTCCGGGAGGCTAGTTACCTTTAAGTTAT 6610
 Qy 771 TGGTATGACTGGTTTTAAGGCAAAAGTGGTTTTTTCGTTACCTTAAATGATGATCGTT 830
 Db 6611 TGGTATGACTGGTTTTAAGGCAAAAGTGGTTTTTTCGTTACCTTAAATGATGATCG-- 6668
 Qy 831 TTAGAAAACCGACTGTAAAAGTACAGTCGGCAATTAATCTCATATTTATAAAGCCAGTCAAT 890
 Db 6669 TTAGAAAACCGACTGTAAAAGTACAGTCGGCAATTAATCTCATATTTATAAAGCCAGTCAAT 6728
 Qy 891 TAGGCTATCTGACAAATCTCTGAATAGAGTTTCAATAACATCTCTGATGATTAACCATCAC 950
 Db 6729 TAGGCTATCTGACAAATCTCTGAATAGAGTTTCAATAACATCTCTGATGATTAACCATCAC 6788
 Qy 951 AAACAGAAATGATGTACCTGTAAAGATAGCGGTAAATATATTGAATTAACCTTTTAAATGA 1010
 Db 6789 AAACAGAAATGATGTACCTGTAAAGATAGCGGTAAATATATTGAATTAACCTTTTAAATGA 6848
 Qy 1011 ATTTTCTCTGCTAATAATAGGTTAGAGTAAATTAATTAATTAATTAATTAATTAATTAATTA 1070
 Db 6849 ATTTTCTCTGCTAATAATAGGTTAGAGTAAATTAATTAATTAATTAATTAATTAATTAATTA 6908
 Qy 1071 ACCCAGTAATGAAATGCTGCAATTAAGAGGAAAGAGCAATTTTCAGTATAGGTG 1130
 Db 6909 ACCCAGTAATGAAATGCTGCAATTAAGAGGAAAGAGCAATTTTCAGTATAGGTG 6968
 Qy 1131 TTTTGGGAAAACAAATTTCCCGAAACCAATTAATTTCTCTACATCAGAAAGGTATTAATCAAT 1190
 Db 6969 TTTTGGGAAAACAAATTTCCCGAAACCAATTAATTTCTCTACATCAGAAAGGTATTAATCAAT 7028
 Qy 1191 AAACTCTTTGAAATGCTTTTACAGAGTCCAAATACAGAGAAATGTTTTAGATACAC 1250
 Db 7029 AAACTCTTTGAAATGCTTTTACAGAGTCCAAATACAGAGAAATGTTTTAGATACAC 7088
 Qy 1251 CATCAAAATTTGTATAAAGTGGCTCTAACTTATCCCAATAACCTAACTCTCCGTCGCTAT 1310
 Db 7089 CATCAAAATTTGTATAAAGTGGCTCTAACTTATCCCAATAACCTAACTCTCCGTCGCTAT 7148
 Qy 1311 TGTAACCAAGTCTAAAAGCTGATTTGATGTTTATCACCCCTGTCTAAGAAAATTAATG 1370
 Db 7149 TGTAACCAAGTCTAAAAGCTGATTTGATGTTTATCACCCCTGTCTAAGAAAATTAATG 7208
 Qy 1371 CAGGTAATAATTTATATCT 1430
 Db 7209 CAGGTAATAATTTATATCT 7268
 Qy 1431 CTGTGGTTATCTAAAAGTCTGTTTGTGGTTCCAAATAATGAATTAATAATCTCTTTCTCT 1490
 Db 7269 CTGTGGTTATCTAAAAGTCTGTTTGTGGTTCCAAATAATGAATTAATAATCTCTTTCTCT 7328
 Qy 1491 TCCAAATGTCTAAATCAATTTTATTAAGTTCAATTTGATGATGCTTCAATTTTATCT 1550
 Db 7329 TCCAAATGTCTAAATCAATTTTATTAAGTTCAATTTGATGATGCTTCAATTTTATCT 7389
 Qy 1551 AAAGTCAATTTTAGAGGCTTACTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1610
 Db 7389 AAAGTCAATTTTAGAGGCTTACTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7448
 Qy 1611 AGTCAATAATTTAGTAACTAAATATATATTTTAAATAATCCGACTTTATCCAAATTTTC 1670
 Db 7449 AGTCAATAATTTAGTAACTAAATATATATTTTAAATAATCCGACTTTATCCAAATTTTC 7508
 Qy 1671 GTTTGTGAACTAATCGGTGCTTTTGTGGAAT--AAAGACCAATTAATAAAGTGGT 1729
 Db 7509 GTTTGTGAACTAATCGGTGCTTTTGTGGAATTAATAAAGTGGT 7568
 Qy 1730 CTTTGTGTTTTTTTAAAGGATTTAGCGTAGCGAAAAATCCTTTCTTTCTTATCTGGA 1789
 Db 7569 CTTTGTGTTTTTTTAAAGGATTTAGCGTAGCGAAAAATCCTTTCTTTCTTATCTTGA 7628

Qy 1790 TAATAAGGTAACATTTTCCGGGATAGACTGTAAATCTTACGATATAAATCCCTCTTC 1849
 Db 7629 TAATAAGGTAACATTTTCCGGGATAGACTGTAAATCTTACGATATAAATCCCTCTTC 7688
 Qy 1850 ATTTTCTAATGTAAATCTTATACCTTTATTAATTAATTAATTAATTAATTAATTAATTAAT 1909
 Db 7689 ATTTTCTAATGTAAATCTTATACCTTTATTAATTAATTAATTAATTAATTAATTAATTAAT 7748
 Qy 1910 TTTCTTATTAAGCAAAATGGCCGATTTAAGCACACCTTTTATTCGTTAATCGGCATGA 1969
 Db 7749 TTTCTTATTAAGCAAAATGGCCGATTTAAGCACACCTTTTATTCGTTAATCGGCATGA 7808
 Qy 1970 CAGCCATGATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2029
 Db 7809 CAGCCATGATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7868
 Qy 2030 AATTATTAAGGTCATCGTTCAAAATGTTATCGGTTTTTACACATCCCATATATATCCGT 2089
 Db 7869 AATTATTAAGGTCATCGTTCAAAATGTTATCGGTTTTTACACATCCCATATATATCCGT 7928
 Qy 2090 GTTCGTTCTCTGCTCAATCTCCGATTCGAAATTTCTCTAGCGATTTCCAGAAATTTCT 2149
 Db 7929 GTTCGTTCTCTGCTCAATCTCCGATTCGAAATTTCTCTAGCGATTTCCAGAAATTTCT 7988
 Qy 2150 CAGAGTCGAAATTTGACCCAGACATTAAGAACTGGCAAGATGTTGTTGTTGTTGTTGTTGTT 2209
 Db 7989 CAGAGTCGAAATTTGACCCAGACATTAAGAACTGGCAAGATGTTGTTGTTGTTGTTGTTGTT 8048
 Qy 2210 GATCTGATTTGCTTAACTGCTTCAAGTAAAGCGAGCGCTCGTGTATTAACAGATCGGAT 2269
 Db 8049 GATCTGATTTGCTTAACTGCTTCAAGTAAAGCGAGCGCTCGTGTATTAACAGATCGGAT 8108
 Qy 2270 GATCAGAGCCAAATCAACATGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2329
 Db 8109 GATCAGAGCCAAATCAACATGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8168
 Qy 2330 ATGTTGTCGTTGCTGCAACGATTAATTAAGCGATTTGCTGCTGCTGCTGCTGCTGCTGCT 2389
 Db 8169 ATGTTGTCGTTGCTGCAACGATTAATTAAGCGATTTGCTGCTGCTGCTGCTGCTGCTGCT 8228
 Qy 2390 CTACGATAAGGGCAAAATCGCATCTGTTGAAACGTTTTGGGCTTCTACCGATTTAGCAGTTT 2449
 Db 8229 CTACGATAAGGGCAAAATCGCATCTGTTGAAACGTTTTGGGCTTCTACCGATTTAGCAGTTT 8288
 Qy 2450 GATACACTTTCTCTAAGTATCCACCTGATTAATTAAGCGAAATTAAGAGAAATTTGAC 2509
 Db 8289 GATACACTTTCTCTAAGTATCCACCTGATTAATTAAGCGAAATTAAGAGAAATTTGAC 8348
 Qy 2510 CATGTGTAAAGGGCCCAATCTGATTTCCACCTGAGATGATTAATTAAGTATCTTTGCTGCT 2569
 Db 8349 CATGTGTAAAGGGCCCAATCTGATTTCCACCTGAGATGATTAATTAAGTATCTTTGCTGCT 8408
 Qy 2570 TATCAAAATTTCACTTTCCACTCTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2629
 Db 8409 TATCAAAATTTCACTTTCCACTCTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8468
 Qy 2630 CCTCTGTGATGATGACACATCATCTCATATCCGATATAGGCGCCCATGATCTGACGAC 2689
 Db 8469 CCTCTGTGATGATGACACATCATCTCATATCCGATATAGGCGCCCATGATCTGACGAC 8528
 Qy 2690 CAAGAGAGCCATTAACCAACCAATAGCCCTTAACATCATCCCATATTTATCCAAATTTGCTT 2749
 Db 8529 CAAGAGAGCCATTAACCAACCAATAGCCCTTAACATCATCCCATATTTATCCAAATTTGCTT 8588
 Qy 2750 CCTTAATTTCAATGAAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2809
 Db 8589 CCTTAATTTCAATGAAATCTTCT 8648
 Qy 2810 CTATTTCTCATTTCCCTTTTACAGATAATTTTATGATTTGCTTTCTTAAATAAGAAATTTTGA 2869
 Db 8649 CTATTTCTCATTTCCCTTTTACAGATAATTTTATGATTTGCTTTCTTAAATAAGAAATTTTGA 8708

QY 2870 GAGCACCCTTTCTTATTCAGCTATTAACTACCTGCTTCTTAAGCATCCTTCAATCCCTTTT 2829
DB |||||
QY 8709 GAGCACCCTTCTTATTCAGCTATTAACTACCTGCTTCTTAAGCATCCTTCAATCCCTTTT 8768
DB |||||
QY 2930 AATAACAATATTAGCATCTAATCTTCAACAACCTGGCCGCTTCTTGAACCTACTCTTTAA 2989
DB |||||
QY 8769 AATAACAATATTAGCATCTAATCTTCAACAACCTGGCCGCTTCTTGAACCTACTCTTTAA 8828
DB |||||
QY 2990 TAAAAATATTTTCGGTTCCTCAATTCACATGCAATAGAAAAATCCATCTTCATCGG 3049
DB |||||
QY 8829 TAAAAATATTTTCGGTTCCTCAATTCACATGCAATAGAAAAATCCATCTTCATCGG 8888
DB |||||
QY 3050 CTTTTTCGTCACTCTGTATGAATCAATCGCCCTTCTTCGTGTCATCAAGGTTTAAAT 3109
DB |||||
QY 8889 CTTTTTCGTCACTCTGTATGAATCAATCGCCCTTCTTCGTGTCATCAAGGTTTAAAT 8948
DB |||||
QY 3110 TTTTATGATTTCTTTTAAACAAACCAATAGGAGATTAACTTTTACGGTGTAAACCTT 3169
DB |||||
QY 8949 TTTTATGATTTCTTTTAAACAAACCAATAGGAGATTAACTTTTACGGTGTAAACCTT 9008
DB |||||
QY 3170 CTTCCAAATCAGACAAACGTTTCAAATCTTTTCTTCATCATCGTGCATATAAATCCGTTAT 3229
DB |||||
QY 9009 CTTCCAAATCAGACAAACGTTTCAAATCTTTTCTTCATCATCGTGCATATAAATCCGTTAT 9068
DB |||||
QY 3230 CTTTACAGGATTTTTCAGCTTTTTCGTCATTTGCGGATTTGATATCCGATTTATATTAT 3289
DB |||||
QY 9069 CTTTACAGGATTTTTCAGCTTTTTCGTCATTTGCGGATTTGATATCCGATTTATATTAT 9128
DB |||||
QY 3290 TTTTCGGT--ATTTTTATTAACAGCTCTCAAAATCGTTTCTGGAGCTTTTAGCGTTTA 3347
DB |||||
QY 9129 TTTTCGGTTCGATTTTATTAACAGCTCTCAAAATCGTTTCTGGAGCTTTTAGCGTTTA 9188
DB |||||
QY 3348 TTTTCGTTTATTCGCGCATATCGTTTAAACAGCGCTTATTCGAGCGTAAAGCCCTTG 3407
DB |||||
QY 9189 TTTTCGTTTATTCGCGCATATCGTTTAAACAGCGCTTATTCGAGCGTAAAGCCCTTG 9248
DB |||||
QY 3408 AGCGTAGCGT--GCTTTGAGGAGAGATGTTGCTGTTAGATTAATGAAGCGGATGACTGA 3466
DB |||||
QY 9249 AGCGTAGCGGCTTTGAGGAGAGATGTTGCTGTTAGATTAATGAAGCGGATGACTGA 9308
DB |||||
QY 3467 ATGAATTAATAGCGCAGCTCTTCTATTTTCGCTTGGAGAGCTCAAGGGAGTTTGAG 3526
DB |||||
QY 9309 ATGAATTAATAGCGCAGCTCTTCTATTTTCGCTTGGAGAGCTCAAGGGAGTTTGAG 9368
DB |||||
QY 3527 GGAATGAATTCCTCATGCGTTTGAATTTTAAATAATTCCTTGAATTTTCCGAGCGGTA 3586
DB |||||
QY 9369 GGAATGAATTCCTCATGCGTTTGAATTTTAAATAATTCCTTGAATTTTCCGAGCGGTA 9428
DB |||||
QY 3587 GCGCTTGAATTTTGAATAAATTTGGAATTTTGAATAAAT--GGGGGGAAGGAAGC 3645
DB |||||
QY 9429 GCGCTTGAATTTTGAATAAATTTGGAATTTTGAATAAATTTGGGGGGAAGGAAGC 9488
DB |||||
QY 3646 GAAATTTTCCTTCCGTTACTACGACCCCTCAATTAAGTGCAGTGCCTCAATTTTGTGCCAA 3705
DB |||||
QY 9489 GAAATTTTCCTTCCGTTACTACGACCCCTCAATTAAGTGCAGTGCCTCAATTTTGTGCCAA 9548
DB |||||
QY 3706 AACGCTCTATCCCACTGGCTCAAGGCTTGAAGGCTTTTCAATCCGCAACGAATCGCC 3765
DB |||||
QY 9549 AACGCTCTATCCCACTGGCTCAAGGCTTGAAGGCTTTTCAATCCGCAACGAATCGCC 9608
DB |||||
QY 3766 AACGTTTTCGCAACGTTTTTATAAATCTATATTTAAGTAGCTTTTATTTGTTGTTTTAT 3825
DB |||||
QY 9609 AACGTTTTCGCAACGTTTTTATAAATCTATATTTAAGTAGCTTTTATTTGTTGTTTTAT 9668
DB |||||
QY 3826 GATTACAAAGTGATACACTAAATTTTATAAATTTATTTGATTTGAGTTTAAATGTTGA 3885
DB |||||
QY 9669 GATTACAAAGTGATACACTAAATTTTATAAATTTATTTGATTTGAGTTTAAATGTTGA 9728
DB |||||
QY 3886 TTTTCAGAAATCGAAAAAAGAGTTATGATTTCTTCGACAAAGAGCAGATATAAATAATTA 3945
DB |||||
QY 9729 TTTTCAGAAATCGAAAAAAGAGTTATGATTTCTTCGACAAAGAGCAGATATAAATAATTA 9788
DB |||||
QY 3946 CAGATATGGCGAAACAAAAAGGTTTTTCAAAATCTGCGG--TTGCGCGCTTAGCTATAGA 4003

DB 9789 CAGATATGGCGAAACAAAAAGGTTTTTCAAAATCTGCGGTTGCGGCTTAGCTATAGA 9848
QY |||||
QY 4004 AGAATATGCAAGAAAGGAATCAGAAACAAAAAATAAGCGAAAGCTCCGTTTTTTAGAA 4063
DB |||||
QY 9849 AGAATATGCAAGAAAGGAATCAGAAACAAAAAATAAGCGAAAGCTCCGTTTTTTAGAA 9908
DB |||||
QY 4064 GATAAGCTTTTCGCTACTCTGTTTTTGTATAGGTAA--TATATCATGGCTATTAAATATAC 4122
DB |||||
QY 9909 GATACAGTTTTTCGCTACTCTGTTTTTGTATAGGTAAATTTATCATGGCTATTAAAAATAC 9968
DB |||||
QY 4123 TAAAGCTAGAAATTTTGGATTTTATTATATCTCTGACTCAATTCCTAAATGATGGAAGA 4182
DB |||||
QY 9969 TAAAGCTAGAAATTTTGGATTTTATTATATCTCTGACTCAATTCCTAAATGATGGAAGA 10028
DB |||||
QY 4183 AAAATTAGAGTTTTGGCGGTATCTATGCTGTCTAGTCTTTTACAGCATATGGACGAAA 4242
DB |||||
QY 10029 AAAATTAGAGTTTTGGCGGTATCTATGCTGTCTAGTCTTTTACAGCATATGGACGAAA 10088
DB |||||
QY 4243 AAAAGATAAAGATACATGGAATAGTAGTCAATGTTATACGAAATGGAAGCACTATAAAA 4302
DB |||||
QY 10089 AAAAGATAAAGATACATGGAATAGTAGTCAATGTTATACGAAATGGAAGCACTATAAAA 10148
DB |||||
QY 4303 ACCACATATCAGCTTTATATATTTTGGACGAAATCTGTAACTAAGAAAGCTTAGGAA 4362
DB |||||
QY 10149 ACCACATATCAGCTTTATATATTTTGGACGAAATCTGTAACTAAGAAAGCTTAGGAA 10208
DB |||||
QY 4363 CAAGATTAAGCGAAATTTGGGGAATAGTTCAGTTGCTCATGTTGAGATACCTTTATAT 4422
DB |||||
QY 10209 CAAGATTAAGCGAAATTTGGGGAATAGTTCAGTTGCTCATGTTGAGATACCTTTATAT 10268
DB |||||
QY 4423 CAAGGTTTCATATGAATTTTGAATCATCAATCAAAAGGACGCTATTCCTAAGATAAACA 4482
DB |||||
QY 10269 CAAGGTTTCATATGAATTTTGAATCATCAATCAAAAGGACGCTATTCCTAAGATAAACA 10328
DB |||||
QY 4483 TATATACGACAAAGAAATTTTGAACATTAATGATTTTGAATTTTGAATTTTGAATTTTGA 4542
DB |||||
QY 10329 TATATACGACAAAGAAATTTTGAACATTAATGATTTTGAATTTTGAATTTTGAATTTTGA 10388
DB |||||
QY 4543 ACTTGATGAAAGCCAAAAAAGAGAAATTTGAAGAAATTTTACTTTTAGATATAGTGGATGACTA 4602
DB |||||
QY 10389 ACTTGATGAAAGCCAAAAAAGAGAAATTTGAAGAAATTTTACTTTTAGATATAGTGGATGACTA 10448
DB |||||
QY 4603 TAAATTTGTAATACAAAGATTTTAAAGCTTTTATTCGCTTAGGGAGCGGAGTTGG 4662
DB |||||
QY 10449 TAAATTTGTAATACAAAGATTTTAAAGCTTTTATTCGCTTAGGGAGCGGAGTTGG 10508
DB |||||
QY 4663 AATTTTAAATACGAATGATGTAAAGAGATTTTGTTCACAAACTCTAGCGCTTTAGATT 4722
DB |||||
QY 10509 AATTTTAAATACGAATGATGTAAAGAGATTTTGTTCACAAACTCTAGCGCTTTAGATT 10568
DB |||||
QY 4723 ATGTTTGAAGGCAATATCATGTTGGATATAGAGCAAGTTATGCAAGGTTCTTGATGC 4782
DB |||||
QY 10569 ATGTTTGAAGGCAATATCATGTTGGATATAGAGCAAGTTATGCAAGGTTCTTGATGC 10628
DB |||||
QY 4783 TGAAACGGGGGAAATTAATAATGACAAACAAAGAAAGAGTTTATTCCTGAAAAATGAGGAA 4842
DB |||||
QY 10629 TGAAACGGGGGAAATTAATAATGACAAACAAAGAAAGAGTTTATTCCTGAAAAATGAGGAA 10688
DB |||||
QY 4843 TTAATAAAGAAATTAAGGACTTTAAAGAGCGTATTTGAAGATACAGAGAAATGGAAGTT 4902
DB |||||
QY 10689 TTAATAAAGAAATTAAGGACTTTAAAGAGCGTATTTGAAGATACAGAGAAATGGAAGTT 10748
DB |||||
QY 4903 GAATTAAGCTACAACTAGATTTATTTGAGGAGGATTTTGAATATAAATAAGCCCCC 4962
DB |||||
QY 10749 GAATTAAGCTACAACTAGATTTATTTGAGGAGGATTTTGAATATAAATAAGCCCCC 10808
DB |||||
QY 4963 CTGACGAAGTCGCGACTTCCTTCTTTT 4992
DB |||||
QY 10809 CTGACGAAGTCGAAGGGGCTTTTATTTT 10838
DB |||||

ID	ACD13842 standard; DNA; 10929 BP.	XX
ACD13842		XX
AC	ACD13842;	XX
AC	ACD13842;	XX
XX	15-AUG-2003 (first entry)	XX
DT		XX
DE	Plasmid pEP25 containing L. lactis CP25 promoter.	XX
XX	Promoter; ds; gram positive bacteria; Staphylococcus aureus; plasmid;	XX
KW	Enterococcus faecalis; operator; xylO; tetO; trpO; malO; lambda-clo;	XX
KW	cellular proliferation.	XX
KW		XX
OS	Lactococcus lactis.	XX
OS	Synthetic.	XX
XX		XX
XX	US2003027286-A1.	XX
PN		XX
XX	06-FEB-2003.	XX
PD		XX
XX	21-DEC-2001; 2001US-00032393.	XX
PF		XX
XX	06-SEP-2000; 2000US-0230335P.	XX
PR	27-DEC-2000; 2000US-0259434P.	XX
XX		XX
PA	(HASE/) HASELBECK R.	XX
PA	(WALL/) WALL D.	XX
PA	(GROS/) GROSS M.	XX
XX		XX
XX	Haselbeck R, Wall D, Gross M;	XX
PI		XX
PI	WPI; 2003-479541/45.	XX
DR		XX
XX	New isolated nucleic acid comprising a fusion promoter having at least	XX
PT	one promoter that is modified to have altered activity in at least one	XX
PT	gram-positive organism, useful for regulating gene expression in	XX
PT	bacteria.	XX
XX		XX
XX	Example 1; Page 58-62; 142pp; English.	XX
PS		XX
CC	The invention relates to an isolated nucleic acid comprising a fusion	XX
CC	promoter having at least one promoter that is modified to have altered	XX
CC	activity in at least one gram-positive organism (e.g. Staphylococcus	XX
CC	aureus or Enterococcus faecalis). The promoter is linked to at least one	XX
CC	operator selected from xylO, tetO, trpO, malO and lambda-clo, which are	XX
CC	positioned such that the binding of at least one repressor to the	XX
CC	operator represses transcription from the fusion promoter. Also included	XX
CC	are a vector comprising the isolated nucleic acid, a host cell comprising	XX
CC	the nucleic acid. The fusion promoter is useful for identifying genes	XX
CC	involved in cellular proliferation, identifying a compound that reduces	XX
CC	the activity or level of a gene product required for proliferation of a	XX
CC	cell, inhibiting the activity or expression of a gene in an operon	XX
CC	required for proliferation, manufacturing an antibiotic, identifying a	XX
CC	gene that is required for proliferation of a prokaryotic cell,	XX
CC	identifying a compound that inhibits the proliferation of a prokaryotic	XX
CC	cell and regulating gene expression in bacteria. The present sequence is	XX
CC	a plasmid containing a fusion promoter of the invention	XX
XX		XX
XX	Sequence 10929 BP; 3259 A; 2175 C; 2157 G; 3338 T; 0 U; 0 Other;	XX
SQL		XX
Query Match	82.3%; Score 4232.8; DB 8; Length 10929;	
Best Local Similarity	99.4%; Pred. No. 0;	
Matches 4323; Conservative	0; Mismatches 17; Indels 10; Gaps 7;	
QY	651 CTGATTTCATTTTCATCTACAAACTGCATACTCATATGTAATAATCGTCTCTTTTA 710	
DB	6491 CTGATTTCATTTTCATCTACAAACTGCATACTCATATGTAATAATCGTCTCTTTTA 6550	
QY	711 GGTGGCACAAATGTGAGGCATTTTCGCTCTTTCCGGCGAGGCTAGTTACCCCTTAAGTTAT 770	
DB	6551 GGTGGCACAAATGTGAGGCATTTTCGCTCTTTCCGGCGAGGCTAGTTACCCCTTAAGTTAT 6610	
QY	771 TGTATGACTGGTTTTTAAAGCGAAAAAGTTGCTTTTTTCGTACCTATTAATATGATCGTT 830	

Qy 4064 GATACGAGTTTTCCTACTTGTGTTTTCGATAGGTAA-TATATCATGCTATTTAAATAATAC 4122
Db 9909 GATACGAGTTTTCCTACTTGTGTTTTCGATAGGTAA-TATATCATGCTATTTAAATAATAC 9968
Qy 4123 TAAGCTAGAAATTTTCGATTTTATTTATATCTGCTCAATCCCTAAATGATGGAAGA 4182
Db 9969 TAAGCTAGAAATTTTCGATTTTATTTATATCTGCTCAATCCCTAAATGATGGAAGA 10028
Qy 4183 AAAATTAGAGAGTTTGGCGGTATCTATGCTGTCAGTCCCTTTACACGATATGACGAAAA 4242
Db 10029 AAAATTAGAGAGTTTGGCGGTATCTATGCTGTCAGTCCCTTTACACGATATGACGAAAA 10088
Qy 4243 AAAAGATAAGATACATGGAATAGTAGTGTATATGGAATGGAAGCACTATAAAA 4302
Db 10089 AAAAGATAAGATACATGGAATAGTAGTGTATATGGAATGGAAGCACTATAAAA 10148
Qy 4303 ACCACATATCACGTTATATATATGACGAAATCCCTGTAACATAGAAAGCGTTAGAA 4362
Db 10149 ACCACATATCACGTTATATATATGACGAAATCCCTGTAACATAGAAAGCGTTAGAA 10208
Qy 4363 CAAGATTAAAGCGAAATTTGGGGAATAGTTTCAAGTTCAGTTCCTGATGATGATGATATAT 4422
Db 10209 CAAGATTAAAGCGAAATTTGGGGAATAGTTTCAAGTTCAGTTCCTGATGATGATATAT 10268
Qy 4423 CAAGATTAAAGCGAAATTTGGGGAATAGTTTCAAGTTCAGTTCCTGATGATGATATAT 4482
Db 10269 CAAGATTAAAGCGAAATTTGGGGAATAGTTTCAAGTTCAGTTCCTGATGATGATATAT 10328
Qy 4483 TATATACGACAAAAAAGATATTTTGAACATTAATGATTTGATATGACCGCTATATAAC 4542
Db 10329 TATATACGACAAAAAAGATATTTTGAACATTAATGATTTGATATGACCGCTATATAAC 10388
Qy 4543 ACTTGATGAAAGCCAAAGAAAGAGATTTGAAGATTTTCTTTTATATAGTGGATGACTA 4602
Db 10389 ACTTGATGAAAGCCAAAGAAAGAGATTTGAAGATTTTCTTTTATATAGTGGATGACTA 10448
Qy 4603 TAAATTGCTAAATACAAAGATTTTAAATGCTTTTATTCGCTTGGGAGCGAGTTTGG 4662
Db 10449 TAAATTGCTAAATACAAAGATTTTAAATGCTTTTATTCGCTTGGGAGCGAGTTTGG 10508
Qy 4663 AATTTTAAATACGAATGATTAAGATATTTGTTTCAACAACTCTAGCGCTTTAGATT 4722
Db 10509 AATTTTAAATACGAATGATTAAGATATTTGTTTCAACAACTCTAGCGCTTTAGATT 10568
Qy 4723 ATGTTTTCAGGCGCAATTCAGTGTGGATATAGAGCAAGTTTATGCAAGGTTCTTTGATGC 4782
Db 10569 ATGTTTTCAGGCGCAATTCAGTGTGGATATAGAGCAAGTTTATGCAAGGTTCTTTGATGC 10628
Qy 4783 TGAACGCGGGGAAATATAATGACAAACAAAGAAAGATTTATTCGTGAAATGAGGAA 4842
Db 10629 TGAACGCGGGGAAATATAATGACAAACAAAGAAAGATTTATTCGTGAAATGAGGAA 10688
Qy 4843 TTAATAAAGAAATTAAGGCTTAAAGAGCGTATTGAAAGATACAGAGAATGGAAGTT 4902
Db 10689 TTAATAAAGAAATTAAGGCTTAAAGAGCGTATTGAAAGATACAGAGAATGGAAGTT 10748
Qy 4903 GAATTAAGTACAAATAGATTATTTAGAGAGGAGGATTTTGAATAATATAAGCCCC 4962
Db 10749 GAATTAAGTACAAATAGATTATTTAGAGAGGAGGATTTTGAATAATATAAGCCCC 10808
Qy 4963 CTGACGAAAGTCGCACTTCGTTCTTTT 4992
Db 10809 CTGACGAAAGTCGCAAGGGGCTTTTATTTT 10838

RESULT 4

AAT33477/c

ID AAT33477 standard; DNA; 8119 BP.

XX

AC AAT33477;

XX

DT 16-DEC-1996 (first entry)

XX

DE Subtilisin N62D/G166D double mutant DNA in phagemid pSS5.
XX Subtilisin BPN'; protease; enzyme engineering; protein engineering;
KW Bacillus amyloliquefaciens; mutagenesis; substrate specificity; ss.
XX Synthetic.
OS
XX
FH Key Location/Qualifiers
FT misc_difference 7 /*tag= a
FT /note= "base n at position 7 is not identified in the
FT specification"
FT CDS 103..1251
FT /*tag= b
FT sig_peptide 103..192
FT /*tag= c
FT mat_peptide 424..1248
FT /*tag= d
FT misc_difference 1971 /*tag= e
FT /note= "base n at position 1971 is not identified in the
FT specification"
FT misc_difference 1985 /*tag= f
FT /note= "base n at position 1985 is not identified in the
FT specification"
FT misc_difference 8119 /*tag= g
FT /note= "base n at position 8119 is not identified in the
FT specification"
XX WO9627671-A1.
XX 12-SEP-1996; 96WO-US002861.
XX 27-FEB-1996; 96WO-US002861.
XX 03-MAR-1995; 95US-00398028.
XX 19-JUL-1995; 95US-00504265.
XX (GETH) GENENTECH INC.
XX Ballinger MD, Wells JA;
XX WPI; 1996-425431/42.
XX P-PSDB; AAW00247.
XX Subtilisin variants for cleaving substrates contg. basic residues - allow
XX effective cleavage of fusion proteins with basic linker sequences.
XX Disclosure; Page 29-35; 83pp; English.
XX A DNA sequence (AAT33476) in phagemid pSS5 codes for the prepro form
XX (AAW00247) of subtilisin N62D/G166D double mutant. The amino acid
XX substitutions alter the substrate specificity of the enzyme, such that it
XX specifically cleaves protein substrates contg. basic amino acid residues
XX at positions P1 and P2. Such an enzyme is useful for cleaving fusion
XX proteins with basic substrate linkers, esp. fusion proteins contg. a
XX dibasic sequence, and for processing hormones or other proteins in vitro
XX or in vivo that contain basic cleavage sites. The DNA sequence can be
XX obtd. by site-directed mutagenesis of the Bacillus amyloliquefaciens
XX subtilisin BPN' gene, and used to produce the variant protease in
XX transformed host cells
XX Sequence 8119 BP; 2459 A; 1628 C; 1811 G; 2217 T; 0 U; 4 Other;
SQ

Query Match

Best Local Similarity 48.0%; Score 2467.2; DB 2; Length 8119;

Matches 2472; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 831 TTGAAAAACCGACTGTAAAAAGTACAGTCGGCGATATCTCATATTATAAAGCCATCAT 890

Db 7847 TTGAAAAACCGACTGTAAAAAGTACAGTCGGCGATATCTCATATTATAAAGCCATCAT 7788

QY	891	TAGGCTATCTGACAAATTCCTGTAATAGAGTTTCATAAAACAATCTCGCATGATAACCATCAC	950	QY	1971	AGCCATGATAAATTAATAATACTAGGAGAAGTTAATAAATAGTAACCAACATGATTAAACA	2030
DB	7787	TAGGCTATCTGACAAATTCCTGTAATAGAGTTTCATAAAACAATCTCGCATGATAACCATCAC	7728	DB	6707	AGCCATGATAAATTAATAATACTAGGAGAAGTTAATAAATAGTAACCAACATGATTAAACA	6648
QY	951	AAACAGAAATGATGACCTGTAAGATAGCGGTAAATATATGAAATACCTTTTAATGA	1010	QY	2031	ATTATAGAGGTGATCGTTCAAAATGGTATGCGTTTTCAGACATCCACTATATATCCGTG	2090
DB	7727	AAACAGAAATGATGACCTGTAAGATAGCGGTAAATATATGAAATACCTTTTAATGA	7668	DB	6647	ATTATAGAGGTGATCGTTCAAAATGGTATGCGTTTTCAGACATCCACTATATATCCGTG	6588
QY	1011	ATTTTCCTGCTGTAATAATAGGTGAGAGGTAAATTAATTAATTAATTAATTAATTAATTA	1070	QY	2091	TGCTTCTGTCCTCACTCCTCAATCCCAATCCAGAAATTTCTTAGCGATTCAGAAATTTCTC	2150
DB	7667	ATTTTCCTGCTGTAATAATAGGTGAGAGGTAAATTAATTAATTAATTAATTAATTAATTA	7608	DB	6587	TGCTTCTGTCCTCACTCCTCAATCCCAATCCAGAAATTTCTTAGCGATTCAGAAATTTCTC	6528
QY	1071	ACCCAGTAAATGAAATCCATCGAATTAATAGAAAGAGAAAGCAATTTTCAGGTATAGGTG	1130	QY	2151	AGAGTCGGAAGTTGACCCAGACATTAAGAACTGGCAGAGATGGTCATTAACCTGAAGGAAG	2210
DB	7607	ACCCAGTAAATGAAATCCATCGAATTAATAGAAAGAGAAAGCAATTTTCAGGTATAGGTG	7548	DB	6527	AGAGTCGGAAGTTGACCCAGACATTAAGAACTGGCAGAGATGGTCATTAACCTGAAGGAAG	6468
QY	1131	TTTTGGGAAACAAATTTCCCGAACCATTAATTAATTTCTTACATCAGAAAGGTATTAATCAT	1190	QY	2211	ATCTGATTTGCTTAATCTGCTTCACTTAAGACCGAAGCGCTCGTCTGATTAACAGATGCGATG	2270
DB	7547	TTTTGGGAAACAAATTTCCCGAACCATTAATTAATTTCTTACATCAGAAAGGTATTAATCAT	7488	DB	6467	ATCTGATTTGCTTAATCTGCTTCACTTAAGACCGAAGCGCTCGTCTGATTAACAGATGCGATG	6408
QY	1191	AAAACCTTTTGAAGTCATTTCTTACAGAGTCCAAATACAGAGAAATGTTTTAGATACAC	1250	QY	2271	ATGCGAGCAAAATCAACATGGCACCTGCGCATTTGCTAAGCTGTACAGTCAAGGATGGTAGAAA	2330
DB	7487	AAAACCTTTTGAAGTCATTTCTTACAGAGTCCAAATACAGAGAAATGTTTTAGATACAC	7428	DB	6407	ATGCGAGCAAAATCAACATGGCACCTGCGCATTTGCTAAGCTGTACAGTCAAGGATGGTAGAAA	6348
QY	1251	CATCAAAAATTTGATTAAGTGGCTCTAACTTATCCCAATTAACCTTAACTCTCCGTCGCTAT	1310	QY	2331	TGTTCTGCGTCTTTCGACACGCAATTAATTAAGCCATTTGCGCTTCTACCGATTTAGCAGTTTG	2450
DB	7427	CATCAAAAATTTGATTAAGTGGCTCTAACTTATCCCAATTAACCTTAACTCTCCGTCGCTAT	7368	DB	6287	TGTTCTGCGTCTTTCGACACGCAATTAATTAAGCCATTTGCGCTTCTACCGATTTAGCAGTTTG	6228
QY	1311	TGTAACAGATTTTAAAGCTGATTTGAGTTTATCACCTTGTCACATAAGAAATTAATG	1370	QY	2451	ATPACATTTTCTTAAGTATCCACCTGAATCATTAATTCGGGAAATAGAGAAATTTGACCC	2510
DB	7367	TGTAACAGATTTTAAAGCTGATTTGAGTTTATCACCTTGTCACATAAGAAATTAATG	7308	DB	6227	ATPACATTTTCTTAAGTATCCACCTGAATCATTAATTCGGGAAATAGAGAAATTTGACCC	6168
QY	1371	CAGGTAAATTTTATATCTTTCTTTTATGTTTTGTTTTCGGTATTAACCACTAATATCAATTT	1430	QY	2511	ATGCTGTAAGCGGCCAATCTGATTTCCACCTGAGATGTCATTAATCTAGTAGAATCTCTCGCT	2570
DB	7307	CAGGTAAATTTTATATCTTTCTTTTATGTTTTGTTTTCGGTATTAACCACTAATATCAATTT	7248	DB	6167	ATGCTGTAAGCGGCCAATCTGATTTCCACCTGAGATGTCATTAATCTAGTAGAATCTCTCGCT	6108
QY	1431	CTGTGGTTTATTAAGTCTGTTTGGTTTCAATTAATGATTAATATCTCTTTCTCT	1490	QY	2571	ATCAAAAATCTCTCCACTCCACTCCCGGTGTCATTCATTCAGGTGAACTCTGCTTC	2630
DB	7247	CTGTGGTTTATTAAGTCTGTTTGGTTTCAATTAATGATTAATATCTCTTTCTCT	7188	DB	6107	ATCAAAAATCTCTCCACTCCACTCCCGGTGTCATTCATTCAGGTGAACTCTGCTTC	6048
QY	1491	TCCAAATGCTTAATCAATTTTATTAAGTTTCAATTTGATATGCTCCTCAATTTTATCT	1550	QY	2631	CTCTGTCGACATGACACATCATCTCAATATCCGAATAGGCGCCATCAGTCTGACGACC	2690
DB	7187	TCCAAATGCTTAATCAATTTTATTAAGTTTCAATTTGATATGCTCCTCAATTTTATCT	7128	DB	6047	CTCTGTCGACATGACACATCATCTCAATATCCGAATAGGCGCCATCAGTCTGACGACC	5988
QY	1551	AAAGTGAATTTAGGAGGCTTACTGTCGCTCTCTTCAATTAAGATCAATCTCTTTTAAA	1610	QY	2691	AAAGAGGCCATTAACCAACCAATAGCCTTAACATCCCATATTTTATCCAATATTCGTTTC	2750
DB	7127	AAAGTGAATTTAGGAGGCTTACTGTCGCTCTCTTCAATTAAGATCAATCTCTTTTAAA	7068	DB	5987	AAAGAGGCCATTAACCAACCAATAGCCTTAACATCCCATATTTTATCCAATATTCGTTTC	5928
QY	1611	AGTCAATTAATCTGTAACATAAATATATATTTTAAATAATCCACATTTTATCCAAATTTTC	1670	QY	2751	CTTAATTTATGAGAACAACTTTTCAATCTTCTCTCTAGTCAATTAATTTTGTCCCATTCAC	2810
DB	7067	AGTCAATTAATCTGTAACATAAATATATATTTTAAATAATCCACATTTTATCCAAATTTTC	7008	DB	5927	CTTAATTTATGAGAACAACTTTTCAATCTTCTCTCTAGTCAATTAATTTTGTCCCATTCAC	5868
QY	1671	GTTTGTGAACATTAATGCTTTAGTTGAAGATTAAGAACCAATTAAGAAATGTTGGTC	1730	QY	2811	TATTTCTCAATCTCTTTTTCAGATAATTTTATAGATTTTGTCTTTTCTAAATTAAGAAATTTGGAG	2870
DB	7007	GTTTGTGAACATTAATGCTTTAGTTGAAGATTAAGAACCAATTAAGAAATGTTGGTC	6948	DB	5867	TATTTCTCAATCTCTTTTTCAGATAATTTTATAGATTTTGTCTTTTCTAAATTAAGAAATTTGGAG	5808
QY	1731	TTTTTGTGTTTTTAAAGGATTTGAGGCTAGCGGAAATCTTTTCTTTCTTATCTTGAT	1790	QY	2871	AGCACGGTTTCTTATTCAGCTATTAATTAATCTGCTCTTCTCTAGTCAATTAATTTTGTCCATTTTA	2930
DB	6947	TTTTTGTGTTTTTAAAGGATTTGAGGCTAGCGGAAATCTTTTCTTTCTTATCTTGAT	6888	DB	5807	AGCACGGTTTCTTATTCAGCTATTAATTAATCTGCTCTTCTCTAGTCAATTAATTTTGTCCATTTTA	5748
QY	1791	AATAAGGGTAACTATGTCGGGATAGACTGTAAATTTCTCAGCATTAATAATCCCTTTTCA	1850	QY	2931	ATAACAAATTAAGCATCTAACTTCAACAAACTGGCCCGTTTGTGAACTACTCTTTAAT	2990
DB	6887	AATAAGGGTAACTATGTCGGGATAGACTGTAAATTTCTCAGCATTAATAATCCCTTTTCA	6828	DB	5747	ATAACAAATTAAGCATCTAACTTCAACAAACTGGCCCGTTTGTGAACTACTCTTTAAT	5688
QY	1851	TTTTTCTAATGTAATCTATTAACCTTATTAATTAATTAATTAATTAATTAATTAATTAAT	1910	QY	2991	AAAAATAATTTTTCGTTTCCCAATTCACATTTGCAATTAATAGAAATTCATCTTCATCGGC	3050
DB	6827	TTTTTCTAATGTAATCTATTAACCTTATTAATTAATTAATTAATTAATTAATTAATTAAT	6768	DB	5687	AAAAATAATTTTTCGTTTCCCAATTCACATTTGCAATTAATAGAAATTCATCTTCATCGGC	5628
QY	1911	TCTTATTAACGCAAAATGCGCGATTTAAGCACACCCCTTTATTCGGTTAAATGCGCCATGAC	1970	QY	3051	TTTTTCTGTCATCATCTGTATGAATCAAAATCGCCTTCTTCTGTGTCATCAAGGTTTAAATTT	3110
DB	6767	TCTTATTAACGCAAAATGCGCGATTTAAGCACACCCCTTTATTCGGTTAAATGCGCCATGAC	6708				


```

Db      5627  |||||CGTCATCATCTGATGAATCAATCGCTTCTCTGTGTCATCAAGGTTAATT 5568
Qy      3111  TTATGTATTTCTTTTAAACCAACCATAGAGATTAACCTTTTACGGTGTAAACCTTC 3170
Db      5567  TTATGTATTTCTTTTAAACCAACCATAGAGATTAACCTTTTACGGTGTAAACCTTC 5508
Qy      3171  CTCCAAATCAGACAAACGTTTCAAAATCTTTTCTTCATCATCGGTCAATAAATCGGTATC 3230
Db      5507  CTCCAAATCAGACAAACGTTTCAAAATCTTTTCTTCATCATCGGTCAATAAATCGGTATC 5448
Qy      3231  CTTTACAGATATTTTGCAGTTTCTCAATGCGGATGTATATCCGATTTATATTATT 3290
Db      5447  CTTTACAGATATTTTGCAGTTTCTCAATGCGGATGTATATCCGATTTATATTATT 5388
Qy      3291  TTTTCGGTATTTTATTAAA 3310
Db      5387  TTTTCGGTGAATCAATTGAA 5368

RESULT 5
AAN70271/c
ID AAN70271 standard; DNA; 2059 BP.
XX
AC AAN70271;
XX
DT 01-JAN-1980 (first entry)
XX
DE DNA fragment for transformation of lactic acid bacteria.
XX
KW Lactic acid bacterium; antibiotic-resistance gene; food; cheese; dairy;
XX plasmid; vector; ss.
XX
OS Synthetic.
XX
PN EP228726-A.
XX
PD 15-JUL-1987.
XX
PF 20-NOV-1986; 86EP-00202061.
XX
PR 29-NOV-1985; 85NL-00003316.
XX
PA (NEZU-) NED INST ZUIVEL.
XX
PS WPI; 1987-193108/28.
XX
PT DNA for transforming lactic acid bacteria - comprising partial fragment
PT which codes for a replicon activity in lactic acid bacteria and an
PT antibiotic-resistance gene.
XX
PS Disclosure; Fig 2; 23pp; English.
XX
CC This sequence contains DNA encoding a replicon activity in lactic acid
CC bacteria and at least 1 antibiotic-resistance gene; either kanamycin-
CC resistance and/or chloramphenicol-resistance. Application of this
CC sequence and a plasmid containing it may lead to the development of
CC lactic acid bacteria which produce polypeptides in a stable and efficient
CC manner. The produced lactic acid bacteria would have application in the
CC food industry. See also AAN70272 and AAN70189
XX
SQ Sequence 2059 BP; 672 A; 424 C; 263 G; 700 T; 0 U; 0 Other;

Query Match 32.1%; Score 1650.4; DB 1; Length 2059;
Best Local Similarity 99.5%; Pred. No. 1.2e-287;
Matches 1677; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

Qy      3291  TTTTCGGTATTTTATTAAACGTTCTCAAAATCGTTTCTGGGACGTTTTCAGCGTTTATT 3350
Db      1685  TTTATCGAATTTTATTAAACGTTCTCAAAATCGTTTCTGGGACGTTTTCAGCGTTTATT 1626
Qy      3351  CGTTTAGTATCGGATATCTTTAAACAGCGGTTATCGTAGCGTTAAAGCCCTTGAGC 3410

```

```

Db      1625  CGTTTAGTATCGGATATCTGTTAAACAGCGGTTTATCGTAGCTAAAGCCCTTGAGC 1566
Qy      3411  GTAGCGTCTTTCACGGAAGATGTTGTCTGTAGATTATGAAGCCGATGACTGAATGA 3470
Db      1565  GTAGCGTCTTTCACGGAAGATGTTGTCTGTAGATTATGAAGCCGATGACTGAATGA 1506
Qy      3471  AATAATAAGCGCGACGCTCTCTTCTATTTCGGTTGCGAGGAGCTCAAGGGAGTTTGA 3530
Db      1505  AATAATAAGCGCGACGCTCTCTTCTATTTCGGTTGCGAGGAGCTCAAGGGAGTTTGA 1446
Qy      3531  TGAATTCCTCTCATGGGTTTGATTTTAAAAATTCCTTGCMAATTTTCCGACGGGTAG 3590
Db      1445  TGAATTCCTCTCATGGGTTTGATTTTAAAAATTCCTTGCMAATTTTCCGACGGGTAG 1386
Qy      3591  TGG-AAAAATTTTGAAGAAAAATTTGCAATTTGGAATAATGCGGGAAGGAGCGAAT 3649
Db      1385  TGGAAAAATTTTGAAGAAAAATTTGCAATTTGGAATAATGCGGGAAGGAGCGAAT 1326
Qy      3650  TTTGCTTCCTACTACGACCCCCCATTAAGTCCGAGTCCCAATTTTGTGCCAAAACG 3709
Db      1325  TTTGCTTCCTACTACGACCCCCCATTAAGTCCGAGTCCCAATTTTGTGCCAAAACG 1266
Qy      3710  CTCTATCCCACTGGCTCAAGGGTTTGAGGGGTTTTCATCGCCCAACGAATGCCAACG 3769
Db      1265  CTCTATCCCACTGGCTCAAGGGTTTGAGGGGTTTTCATCGCCCAACGAATGCCAACG 1206
Qy      3770  TTTTCGCCAACGTTTTTATAAATCTATATTTAAGTAGCTTTTATCTGTTTGTATGAT 3829
Db      1205  TTTTCGCCAACGTTTTTATAAATCTATATTTAAGTAGCTTTTATCTGTTTGTATGAT 1146
Qy      3830  ACAAAGTGATACACTAATTTTATAAATTTATGATGGAGTTTTTTAAATCGTGATTC 3889
Db      1145  ACAAAGTGATACACTAATTTTATAAATTTATGATGGAGTTTTTTAAATCGTGATTC 1086
Qy      3890  AGAATCGAAAAAAGAGTTATGATTTCTCTGACAAAAGAGCAAGATAAAAAATTAACAG 3949
Db      1085  AGAATCGAAAAAAGAGTTATGATTTCTCTGACAAAAGAGCAAGATAAAAAATTAACAG 1026
Qy      3950  TATGGCGAAAAAAGAGTTTTCAAAAATCTGCGGTTGCGGCTTAGCTATAGAGAATA 4009
Db      1025  TATGGCGAAAAAAGAGTTTTCAAAAATCTGCGGTTGCGGCTTAGCTATAGAGAATA 966
Qy      4010  TGAAGAAAGGAATCAGAACAAAAAATAAGCGAAAGCTCGCGTTTGTAGAGGATACG 4069
Db      965  TGAAGAAAGGAATCAGAACAAAAAATAAGCGAAAGCTCGCGTTTGTAGAGGATACG 906
Qy      4070  AGTTTCGCTACTGTTTGTATAGTAAGTAATATATCATGCTATTAATAATACTAAAGCT 4129
Db      905  AGTTTCGCTACTGTTTGTATAGTAAGTAATATATCATGCTATT--AAATACTAAAGCT 848
Qy      4130  AGAATTTTGGATTTTATATATCTGACTCAATTTCTTAATGATTTGGAAGAAATTA 4189
Db      847  AGAATTTTGGATTTTATATATCTGACTCAATTTCTTAATGATTTGGAAGAAATTA 788
Qy      4190  GAGAGTTTGGCGCTATCTATGGCTGCAGTCTCTTTACAGATATGACGAAAAAAGAT 4249
Db      787  GAGAGTTTGGCGCTATCTATGGCTGCAGTCTCTTTACAGATATGACGAAAAAAGAT 728
Qy      4250  AAAGATACATGGAATAGTAGTATGATTTACGAAATGGAAGACATATAAAACACAC 4309
Db      727  AAAGATACATGGAATAGTAGTATGATTTACGAAATGGAAGACATATAAAACACAC 668
Qy      4310  TATCAGCTTATATATATGACGAAATCTGTAAACAATAGAAAGCGTTAGGAACAAGAT 4369
Db      667  TATCAGCTTATATATATGACGAAATCTGTAAACAATAGAAAGCGTTAGGAACAAGAT 608
Qy      4370  AAGCGAAAAATGGGAATAGTTGAGTTGCTCATGTTGAGATACCTTGATTTATCAAGGT 4429
Db      607  AAGCGAAAAATGGGAATAGTTGAGTTGCTCATGTTGAGATACCTTGATTTATCAAGGT 548
Qy      4430  TCATATCAATTTGACTCATGATCAAGGAGCGCTATTGCTAAGAAATAACATATATAC 4489
Db      547  TCATATCAATTTGACTCATGATCAAGGAGCGCTATTGCTAAGAAATAACATATATAC 488

```

QY 4490 GACAAAAAGATATTTTGAACATTAATGATTTTGGATTCACCGCTATATAACACATTGAT 4549
 Db 487 GACAAAAAGATATTTTGAACATTAATGATTTTGGATTCACCGCTATATAACACATTGAT 428
 QY 4550 GAAAGCCAAAAGAGATTTGAAGAAATTTACTTTTAGATATAGTGGATGACTATAATTTC 4609
 Db 427 GAAAGCCAAAAGAGATTTGAAGAAATTTACTTTTAGATATAGTGGATGACTATAATTTC 368
 QY 4610 GTAATAACAAAAGATTTAATGGCTTTTATTCGCCCTTAGGGGAGCGGAGTTTGGAAATTTA 4669
 Db 367 GTAATAACAAAAGATTTAATGGCTTTTATTCGCCCTTAGGGGAGCGGAGTTTGGAAATTTA 308
 QY 4670 AATACGAATGATTAAGAATATTTGTTTCAACAACTAGCGCTTTAGATATGTTTT 4729
 Db 307 AATACGAATGATTAAGAATATTTGTTTCAACAACTAGCGCTTTAGATATGTTTT 248
 QY 4730 GAGGCAATATCAGTGTGGATATAGACCAAGTTATGCAAGGTTCTTGATGCTGAAACG 4789
 Db 247 GAGGCAATATCAGTGTGGATATAGACCAAGTTATGCAAGGTTCTTGATGCTGAAACG 188
 QY 4790 GGGGAATTAATAATGACAAACAAAGAAAGATTTATTTGCTGAAATGAGGAATTAATAA 4849
 Db 187 GGGGAATTAATAATGACAAACAAAGAAAGATTTATTTGCTGAAATGAGGAATTAATAA 128
 QY 4850 AAGAAATTAAGGACTTTAAAGAGCGTATTGAAGATACAGAGAATGGAAGTTGAATTA 4909
 Db 127 AAGAAATTAAGGACTTTAAAGAGCGTATTGAAGATACAGAGAATGGAAGTTGAATTA 68
 QY 4910 GTACAACTATAGATTTATGAGAGGAGGATTTATGAAATTAATAAAGCCCGCTGACGA 4969
 Db 67 GTACAACTATAGATTTATGAGAGGAGGATTTATGAAATTAATAAAGCCCGCTGACGA 8
 QY 4970 AAGTCG 4975
 Db 7 AAGTCG 2
 RESULT 5
 AAQ48463
 ID AAQ48463 standard; DNA; 3792 BP.
 AC AAQ48463;
 XX
 XX 25-MAR-2003 (revised)
 DT 18-MAR-1994 (first entry)
 XX
 XX Plasmid pG-host4 containing Ts replication system.
 DE
 XX Temperature sensitive replication; antibiotic resistance marker gene;
 KW site-specific recombination; chromosomal integration; inactivation;
 KW heterologous gene expression; thermosensitive plasmid; ds.
 XX
 OS Synthetic.
 XX
 XX WO9318164-A1.
 PN
 XX
 XX 16-SEP-1993.
 PD
 XX 12-MAR-1993; 93WO-FR000248.
 PF
 XX 13-MAR-1992; 92FR-00003034.
 PR
 XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
 XX
 XX Gruss A, Maguin E;
 PI
 XX
 XX WPI; 1993-303478/38.
 DR
 XX New bacterial plasmid contg. heat sensitive replication system - and
 PT marker gene, opt. capable of chromosomal integration, used to inactivate
 PT specific gene or introduce heterologous gene.
 XX

PS Example 2; Fig 9; 73pp; French.
 XX Plasmid pGK12 (Appl. Environ. Microbiol., 48; 726 (1984)) contg. two
 CC antibiotic resistance marker genes was subjected to mutagenesis with
 CC hydroxylamine. A heat-stable mutant was isolated (coding for a heat-
 CC sensitive RepA - see AAQ48466 and AAQ48467), cut with ClaI and HpaII and
 CC the 3340bp fragment lacking the Cm resistance gene was ligated to a 445bp
 CC PvuII fragment of pBluescript SK+ containing a multicloning site, T7 and
 CC T3 promoters and sequencing primer binding sites. The resulting plasmid
 CC was designated pV6604 (for pG-host4 - AAQ48463). It is thermosensitive in
 CC all hosts tested, including E.coli, and must be maintained at 28 deg.C.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 3792 BP; 1249 A; 552 C; 742 G; 1249 T; 0 U; 0 Other;
 Query Match 31.8%; Score 1636.8; DB 2; Length 3792;
 Best Local Similarity 98.3%; Pred. No. 3.7e-285;
 Matches 1686; Conservative 0; Mismatches 25; Indels 3; Gaps 3;
 QY 3279 TTTATATTTATTTTCGGTATTTTATTAAGAGCTCTCAAAATCGTTTCGGACGTTT 3338
 Db 1181 TATATATATTTATATCGCATTTTATTAAGAGCTCTCAAAATCGTTTCGGACGTTT 1240
 QY 3339 TAGCGTTTATTTTCGTTAGTTATCGGCATAATCGTTTAAACAGCGGTTATCGTAGCGTAA 3398
 Db 1241 TAGCGTTTATTTTCGTTAGTTATCGGCATAATCGTTTAAACAGCGGTTATCGTAGCGTAA 1300
 QY 3399 AAGCCCTTGAGCGGTAGCGT-GCTTTGCAGCGAAGATGTTCTGTTAGATTATGAAAGCC 3457
 Db 1301 AAGCCCTTGAGCGGTAGCGTGGCTTTGCAGCGAAGATGTTCTGTTAGATTATGAAAGCC 1360
 QY 3458 GATGACTGAATGAATAATAAGCGCAGCGCTCTCTATTTTCGGTTGAGGAGGCTCAAGG 3517
 Db 1361 GATGACTGAATGAATAATAAGCGCAGCGCTCTCTATTTTCGGTTGAGGAGGCTCAAGG 1420
 QY 3518 GAGTTTGAGGGAATGAATTCCTCATCGGTTTGATTTTAAATAATGCTGCAATTTTGC 3577
 Db 1421 GAGTATGAGGGAATGAATTCCTCATCGGTTTGATTTTAAATAATGCTGCAATTTTGC 1480
 QY 3578 CGAGCGGTAGCGCTGGAATAATTTTGAATAAATTTTGAATTTGGAATAATGCGGGGA 3637
 Db 1481 CGAGCGGTAGCGCTGGAATAATTTTGAATAAATTTTGAATAATGCGGGGA 1540
 QY 3638 AAGGAAGCGAATTTTTCGTTAGTACGACCCCATTAAGTCCGAGTGCCTATTTT 3697
 Db 1541 AAGGAAGCGAATTTTTCGTTAGTACGACCCCATTAAGTCCGAGTGCCTATTTT 1600
 QY 3698 GTGCCAAAACGCTCTATCCCAACTGGCTCAAGGGTTTTCAGGGTTTTCATATGCCAAC 3757
 Db 1601 GTGCCAAAACGCTCTATCCCAACTGGCTCAAGGGTTTTCAGGGTTTTCATATGCCAAC 1660
 QY 3758 GAATGCCAAAGCTTTTCGCAACGTTTATTAATCTATATTAAGTAGCTTATGTT 3817
 Db 1661 GAATGCCAAAGCTTTTCGCAACGTTTATTAATCTATATTAAGTAGCTTATGTT 1720
 QY 3818 GTTTTATGATTACAAAGTGATACACTAATTTTATAAATAATTTTGAITGGAGTTT 3877
 Db 1721 GTTTTATGATTACAAAGTGATACACTAATTTTATAAATAATTTTGAITGGAGTTT 1780
 QY 3878 AATGGTATTTCAGATCGAAAAAAGAGTTATGATTTCTCGACAAAGAGCAGATAA 3937
 Db 1781 AATGGTATTTCAGATCGAAAAAAGAGTTATGATTTCTCGACAAAGAGCAGATAA 1840
 QY 3938 ABAATTTACAGATATGCGCAACAAAAAGGTTTTCATAAATCTGCGGCTTACG 3997
 Db 1841 ABAATTTACAGATATGCGCAACAAAAAGGTTTTCATAAATCTGCGGCTTACG 1900
 QY 3998 TATAGAAGATATGCAAGAAAGGAATCAGAA CAAAAAATAAAGCGAAGCTCGGTTT 4057
 Db 1901 TATAGAAGATATGCAAGAAAGGAATCAGAA CAAAAAATAAAGCGAAGCTCGGTTT 1960
 QY 4058 TAGAAGGATACGAGTTTTCGCTACTTGTGTTTATGAAGGTAAT-TATATCATGGCTATTAA 4116

1361	DB	TAGAAGGATACGAGGTTTTCGCTACTTGTTTTGATAAGGTAATATATATCAATGCTATTAA	2020
4117	QY	AAATACATAAAGCTAGAAATTTTGGATTTTTTATTATATCTCTGACTCAATTCCTTAATGATTG	4176
2021	DB	AAATACATAAAGCTAGAAATTTTGGATTTTTTATTATATCTCTGACTCAATTCCTTAATGATTG	2080
4177	QY	GAAGAATAAATATAGAGAGTTTGGCGGTATCTATGGCTGTCACTTCCTTTTACAGATATGGA	4236
2081	DB	GAAGAATAAATATAGAGAGTTTGGCGGTATCTATGGCTGTCACTTCCTTTTACAGATATGGA	2140
4237	QY	CGAAAAAAGATAAAGATACATGGAAATAGTAGTGATGTTTATACGAAATCGAAAGACACTA	4296
2141	DB	CGAAAAAAGATAAAGATACATGGAAATAGTAGTGATGTTTATACGAAATCGAAAGACACTA	2200
4297	QY	TAAAAAACCCACTATACGCTTATATATATTTGACGAAATCTGTATACAAATAGAAAGCGT	4356
2201	DB	TAAAAAACCCACTATACGCTTATATATATTTGACGAAATCTGTATACAAATAGAAAGCGT	2260
4357	QY	TAGGAAACAAGATTAAAGCGAAAAATTTGGGGAATAGTTTCAGTTTGCTCATGTTCCAGATACCTTGA	4416
2261	DB	TAGGAAACAAGATTAAAGCGAAAAATTTGGGGAATAGTTTCAGTTTGCTCATGTTCCAGATACCTTGA	2320
4417	QY	TTATATCAAAAGGTTTCAATATGAATATTTGACTCATATGAATCAAAAGACGCTATTGCTAAGAA	4476
2321	DB	TTATATCAAAAGGTTTCAATATGAATATTTGACTCATATGAATCAAAAGACGCTATTGCTAAGAA	2380
4477	QY	TAAACATATATACGACAAAAAAGATATTTTGAACATTAAATGATTTTGATATTCACCGCTA	4536
2381	DB	TAAACATATATACGACAAAAAAGATATTTTGAACATTAAATGATTTTGATATTCACCGCTA	2440
4537	QY	TATAACACTTGATGAAGCCAAAAAGAGAAATGAAGAAATTTACTTTTAGATATAGTGGGA	4596
2441	DB	TATAACACTTGATGAAGCCAAAAAGAGAAATGAAGAAATTTACTTTTAGATATAGTGGGA	2500
4597	QY	TGACTATAATTTGCTTAATACAAAGATTTAAATGGCTTTTATTTCGCTTTAGGGAGCGGA	4656
2501	DB	TGACTATAATTTGCTTAATACAAAGATTTAAATGGCTTTTATTTCGCTTTAGGGAGCGGA	2560
4657	QY	GTTTGGAAATTTTAAATACGAATCATGTATAAAGATATTTGTTTCAACAAACTCTAGCGCCTT	4716
2561	DB	GTTTGGAAATTTTAAATACGAATCATGTATAAAGATATTTGTTTCAACAAACTCTAGCGCCTT	2620
4717	QY	TAGATTTATGCTTTTGAGGCGCAATTTATCAGTGTGGATATAGAGCAAGTTATGCAAGGTTCT	4776
2621	DB	TAGATTTATGCTTTTGAGGCGCAATTTATCAGTGTGGATATAGAGCAAGTTATGCAAGGTTCT	2680
4777	QY	TGATGCTGAAACGGGGGAAATATAAATGACAAAAAAGAAAAAGATTTATTTCGTGAAAT	4836
2681	DB	TGATGCTGAAACGGGGGAAATATAAATGACAAAAAAGAAAAAGATTTATTTCGTGAAAT	2740
4837	QY	GAGGAATTAATAAAGAAATTAAGGACTTAAAGAGCGGTTATGGAAGATACAGAGAAATG	4896
2741	DB	GAGGAATTAATAAAGAAATTAAGGACTTAAAGAGCGGTTATGGAAGATACAGAGAAATG	2800
4897	QY	GAAGTTCGAATTAAGGTACCAAAATAGATTTTATGAGAGGCGGATTTATTGAATAATAATAAA	4956
2801	DB	GAAGTTCGAATTAAGGTACCAAAATAGATTTTATGAGAGGCGGATTTATTGAATAATAATAAA	2860
4957	QY	GCCCCCTGACGAAAGTCGCACTTCGTTCTTTTTT	4992
2861	DB	G-CGCCCTGACGAAAGTCGAGGGGGTTTTTATTTTT	2895

```

RESULT 7
AAQ48464
ID ID AAQ48464 standard; DNA; 5234 bp.
XX XX
AC AC
AC AAQ48464;
XX XX
DT 25-MAR-2003 (revised)
DT DT
DT 18-MAR-1994 (first entry)
XX XX
DE plasmid pG-host5 containing Ts replication system.

```

XX	Temperature sensitive replication; antibiotic resistance marker gene;
KW	site-specific recombination; chromosomal integration; inactivation;
KW	heterologous gene expression; thermosensitive plasmid; ds.
OS	Synthetic.
XX	
PH	Location/Qualifiers
FT	36..1496
FT	/*tag= a
FT	/standard_name= "ORI"
FT	/note= "origin of replication from pBR322"
FT	2640..4383
FT	/*tag= b
FT	/note= "from pGK12 (derived from pWV01)"
FT	4384..4786
FT	/*tag= c
FT	/note= "from pUB110"
FT	4787..5234
FT	/*tag= d
FT	/note= "from psk"
XX	
PN	W09318164-A1.
XX	
PD	16-SEP-1993.
XX	
EF	12-MAR-1993; 93WO-FR000248.
PR	13-MAR-1992; 92FR-00003034.
XX	
PA	(INRG) INRA INST NAT RECH AGRONOMIQUE.
XX	
PI	Gruss A, Maguin E;
XX	
DR	WPI; 1993-303478/38.
XX	
PT	New bacterial plasmid contg. heat sensitive replication system - and
PT	marker gene, opt. capable of chromosomal integration, used to inactivate
PT	specific gene or introduce heterologous gene.
XX	
PS	Example 2; Fig 10; 73pp; French.
XX	
CC	Plasmid pGK12 (Appl.Environ.Microbiol., 48: 726 (1984)) contg: two
CC	antibiotic resistance marker genes was subjected to mutagenesis with
CC	hydroxylamine. A heat-stable mutant was isolated (coding for a heat-
CC	sensitive RspA - see AAQ48466 and AAQ48467), cut with ClaI and HpaII and
CC	the 3340bp fragment lacking the Cm resistance gene was ligated to a 445bp
CC	PvuII fragment of pBluescript SK+ containing a multicloning site, T7 and
CC	T3 promoters and sequencing primer binding sites. The resulting plasmid
CC	was designated pVE6004 (or pG+host4 - AAQ48463). To facilitate cloning in
CC	E. coli, the 1.4kb Avar-A1-WNT fragment of pBR322 (containing the origin of
CC	replication) was inserted into NsiI-cleaved pG+host4 to give pG+host5
CC	(AAQ48464). The pBR322 ORI allows maintenance of the plasmid in E.coli at
CC	37 deg.C, while the heat-sensitive ORI allows maintenance at 28 deg.C in
CC	gram-positive bacteria. (Updated on 25-MAR-2003 to correct PN field.)
XX	
SO	Sequence 5234 BP; 1569 A; 973 C; 1133 G; 1559 T; 0 U; 0 Other;

Query Match	31.8%;	Score 1635.2;	DB 2;	Length 5234;
Best Local Similarity	98.2%;	Pred. No. 7.5e-285;		
Matches 1685; Conservative	0;	Mismatches 28;	Indels 3;	Gaps 3;

Qy	3279	TTTATATTTATTTTCGGTATTTTTTATTAAACGCTCAAATCGTTTCTGGACGTTT	3338
Db	2623	TATATATATTTATATCCGATTTTTTATTAAACGCTCAAATCGTTTCTGAGACGTTT	2682
Qy	3339	TACGGTTATTTTCGTTTAGTTATTCGGCATTAATCGTTAAACAGCGCTTATTCGTAGCGTAA	3398
Db	2683	TAGCGTTATTTTCGTTTAGTTATTCGGCATTAATCGTTAAACAGCGCTTATTCGTAGCGTAA	2742
Qy	3399	AAGCCCTTGAGCGTAGGGT-GCTTTGCGCGGAAGATGTTGTCGTGTAGATTATGAAGCC	3457
Db	2743	AAGCCCTTGAGCGTAGCGTGGCTTTCAGCGGAAGATGTTGTCGTGTAGATTATGAAGCC	2802

3458 GATGACTGAATGAATAAATAAGCGAGCGTCTCTTATTTTCGGTTGGAGGCTCAAGG 3517
Db
2803 GATGACTGAATGAATAAATAAGCGAGCGTCTCTTATTTTCGGTTGGAGGCTCAAGG 2862
Qy GAGTTGAGGGAATGAATAAATAAGCGAGCGTCTCTTATTTTCGGTTGGAGGCTCAAGG 3577
Db GAGTATGAGGGAATGAATAAATAAGCGAGCGTCTCTTATTTTCGGTTGGAGGCTCAAGG 2922
Qy CGAGCGTACGCTGGAATAAATAAGCGAGCGTCTCTTATTTTCGGTTGGAGGCTCAAGG 3637
Db CGAGCGTACGCTGGAATAAATAAGCGAGCGTCTCTTATTTTCGGTTGGAGGCTCAAGG 2982
Qy AAGGAAGCGAATTTGCTTCCGTAATGACGACCGCCCAATTAAGTCCGAGTCCCAATTTT 3697
Db AAGGAAGCGAATTTGCTTCCGTAATGACGACCGCCCAATTAAGTCCGAGTCCCAATTTT 3042
Qy GTGCGAAGCGTCTATCCCACTGCTCAAGGCTTTGAGGCTTTTCAATGCCCAAC 3757
Db GTGCGAAGCGTCTATCCCACTGCTCAAGGCTTTGAGGCTTTTCAATGCCCAAC 3102
Qy GAATCGCAAGCGTCTATCCCACTGCTCAAGGCTTTGAGGCTTTTCAATGCCCAAC 3817
Db GAATCGCAAGCGTCTATCCCACTGCTCAAGGCTTTGAGGCTTTTCAATGCCCAAC 3162
Qy GTTTTATGATTAACAAGTACACCTAATTTTATAAATCTATATTTAAGTAGCTTTATGTT 3877
Db GTTTTATGATTAACAAGTACACCTAATTTTATAAATCTATATTTAAGTAGCTTTATGTT 3222
Qy AATGCTGATTCAGAAATCGAAAGAGCTATGATTTCTCTGACAAAGAGCAAGATAA 3937
Db AATGCTGATTCAGAAATCGAAAGAGCTATGATTTCTCTGACAAAGAGCAAGATAA 3282
Qy AAAATTAACAGATATCGGGAACAAAGAGCTATGATTTCTCTGACAAAGAGCAAGATAA 3997
Db AAAATTAACAGATATCGGGAACAAAGAGCTATGATTTCTCTGACAAAGAGCAAGATAA 3342
Qy TATAGAGATATGCAAGAGAGATCAGAAACAAAGAGCTATGATTTCTCTGACAAAGAGCAAGATAA 4057
Db TATAGAGATATGCAAGAGAGATCAGAAACAAAGAGCTATGATTTCTCTGACAAAGAGCAAGATAA 3402
Qy TAGAGGATACGAGTTTTCGCTACTGTTTGTATGATGATTAATATCATGCTGCTATTA 4116
Db TAGAGGATACGAGTTTTCGCTACTGTTTGTATGATGATTAATATCATGCTGCTATTA 3462
Qy AATATCAAGCTAGAAATTTGGATTTTATATATCTGACTCAATTCCTATGATG 4176
Db AATATCAAGCTAGAAATTTGGATTTTATATATCTGACTCAATTCCTATGATG 3522
Qy GAAGAAATATGAGAGTTTGGGCTATCTATGCTCTGCTCTGCTCTGCTCTGCTCTGCT 4236
Db GAAGAAATATGAGAGTTTGGGCTATCTATGCTCTGCTCTGCTCTGCTCTGCTCTGCT 3582
Qy CGAAAAAAGATATAAGATACATGATGATGATGATGATGATGATGATGATGATGATGAT 4296
Db CGAAAAAAGATATAAGATACATGATGATGATGATGATGATGATGATGATGATGATGAT 3642
Qy TAAAAACCACTATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4356
Db TAAAAACCACTATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3702
Qy TAGGAAACAGATTAAGCGAATTTGGGAAATAGTTTCAGTTCTGCTCATGTTGAGATCTTGA 4416
Db TAGGAAACAGATTAAGCGAATTTGGGAAATAGTTTCAGTTCTGCTCATGTTGAGATCTTGA 3762
Qy TTATATCAAGGTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4476
Db TTATATCAAGGTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3822
Qy TAAACATATATACGACAAAGAGATTTTGAACATTAATGATTTGATATTTGACCGCTA 4536
Db TAAACATATATACGACAAAGAGATTTTGAACATTAATGATTTGATATTTGACCGCTA 3882

Qy 4537 TATACACTTGTATGAAAGCCAAAGAGAAATGGAAGAAATTTACTTTTAGATATAGTGA 4596
Db 3883 TATACACTTGTATGAAAGCCAAAGAGAAATTTACTTTTAGATATAGTGA 3942
Qy 4597 TGACTATATTTGGTAAATACAAAGATTTAATGCTTTTATTCGCTTAGGGAGCGGA 4656
Db 3943 TGACTATATTTGGTAAATACAAAGATTTAATGCTTTTATTCGCTTAGGGAGCGGA 4002
Qy 4657 GTTTCGAATTTTAAATACGAATGATGAAAGATTTGTTTCAACAACTCTAGCGCTT 4716
Db 4003 GTTTCGAATTTTAAATACGAATGATGAAAGATTTGTTTCAACAACTCTAGCGCTT 4062
Qy 4717 TAGATTATGTTTGGAGGCAATTTACAGTGTGGATATAGAGCAAGTTATGCAAGGTTCT 4776
Db 4063 TAGATTATGTTTGGAGGCAATTTACAGTGTGGATATAGAGCAAGTTATGCAAGGTTCT 4122
Qy 4777 TGATCTGAATCGGGGAAATTAATGATGAAAGATTTAATGCTTTTATTCGCTTAGGGAGCGGA 4836
Db 4123 TGATCTGAATCGGGGAAATTAATGATGAAAGATTTAATGCTTTTATTCGCTTAGGGAGCGGA 4182
Qy 4837 GAGGAATTAAGAAAGAAATTAAGGACTTAAAGAGCGTATTTGAAAGATACAGAGAAATG 4896
Db 4183 GAGGAATTAAGAAAGAAATTAAGGACTTAAAGAGCGTATTTGAAAGATACAGAGAAATG 4242
Qy 4897 GAAGTTGAATTAAGTACAAATTAAGGACTTAAAGAGCGTATTTGAAAGATACAGAGAAATG 4956
Db 4243 GAAGTTGAATTAAGTACAAATTAAGGACTTAAAGAGCGTATTTGAAAGATACAGAGAAATG 4302
Qy 4957 GCGCCCTGACGAAAGTCCGCACTTCGTTCTTTT 4992
Db 4303 G-CGCCCTGACGAAAGTCCGCACTTCGTTCTTTT 4337

RESULT 8
AAQ48465
ID AAQ48465 standard; DNA; 6722 BP.
XX AAC48465;
AC
XX 25-MAR-2003 (revised)
DT 18-MAR-1994 (first entry)
XX
XX Plasmid pG-host6 containing Ts replication system.
XX Temperature sensitive replication; antibiotic resistance marker gene;
KW site-specific recombination; chromosomal integration; inactivation;
KW heterologous gene expression; thermosensitive plasmid; ds.
XX
OS Synthetic.
XX
XX WO918164-A1.
PN
XX
XX 16-SEP-1993.
PD
XX 12-MAR-1993; 93WO-FR000248.
PF
XX 13-MAR-1992; 92FR-00003034.
PR
XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
PA
XX Gruss A, Maguin E;
PI
XX WPI; 1993-303478/38.
DR
XX New bacterial plasmid contg. heat sensitive replication system - and
PT marker gene, opt. capable of chromosomal integration, used to inactivate
PT specific gene or introduce heterologous gene.
XX
XX Disclosure; Fig 11; 73pp; French.
XX
XX Plasmid pGK12 (Appl. Environ. Microbiol., 48; 726 (1984)) contg. two
CC antibiotic resistance marker genes was subjected to mutagenesis with
CC hydroxylamine. A heat-stable mutant was isolated (coding for a heat-

sensitive RepA - see AA048466 and AA048467), cut with ClaI and HpaII and the 3340bp fragment lacking the Cm resistance gene was ligated to a 450bp PvuII fragment of pBluescript SK+ containing a multicloning site, 77 and T3 promoters and sequencing primer binding sites. The resulting plasmid was designated pV6004 (for pG+host4 - AA048463). To facilitate cloning in E.coli, the Aval-EcoRI fragment of pBR322 (containing the origin of replication and the ampicillin resistance gene) was inserted into NsiI-cleaved, blunt-ended pG+host4 to give pG+host6 (AA048466). The pBR322 ORI allows maintenance of the plasmid in E.coli at 37 deg.C while the heat-sensitive ORI allows maintenance at 28 deg.C in gram-positive bacteria. (Updated on 25-MAR-2003 to correct PN field.)

Sequence 6722 BP; 1961 A; 1278 C; 1503 G; 1980 T; 0 U; 0 Other;

Query Match 31.8%; Score 1635.2; DB 2; Length 6722;
Best Local Similarity 98.2%; Pred. No. 7.7e-285;
Matches 1685; Conservative 0; Mismatches 28; Indels 3; Gaps 3;

QY	3279	TTTATATTTATTTTCGGTATTTTATTTAAACGCTCAAAATCGTTTCGGGACGTTT	3338
DB	4111	TATATATATTTATTCGATTTTATTTAAACGCTCAAAATCGTTTCGGGACGTTT	4170
QY	3339	TAGCGTTATTTTCGGTATTTTCGCTATCGCATATCGTAAACAGCGTTATCGTAGCGTAA	3398
DB	4171	TAGCGTTATTTTCGGTATTTTCGCTATCGCATATCGTAAACAGCGTTATCGTAGCGTAA	4230
QY	3399	AAGCCCTTGAGCGTAGCGT - GCTTTTCAGCGAAGATGTTGCTGTAGATTATGAAGCC	3457
DB	4231	AAGCCCTTGAGCGTAGCGTGTGCTTCAGCGAAGATGTTGCTGTAGATTATGAAGCC	4290
QY	3458	GATGACTGAATGAATTAAGCGCAGCGTCTCTATTCGGTTGGAGGAGGCTCAAGG	3517
DB	4291	GATGACTGAATGAATTAAGCGCAGCGTCTCTATTCGGTTGGAGGAGGCTCAAGG	4350
QY	3518	GAGTTTGAGGGAATGAATTCCTCATGCGTTTGATTTTAAAAATGCTGTGCAATTTTCG	3577
DB	4351	GAGTATGAGGGAATGAATTCCTCATGCGTTTGATTTTAAAAATGCTGTGCAATTTTCG	4410
QY	3578	CGAGCGGTAGCGTGAATAATTTTGAATAAATTTGGAATTTGGAATAAATGCGGGA	3637
DB	4411	CGAGCGGTAGCGTGAATAATTTTGAATAAATTTGGAATTTGGAATAAATGCGGGA	4470
QY	3638	AAGGAAGCGAATTTTTCCTTCGGTCTAGCAGCCCGATTAAGTCCGAGTGCCTATTTT	3697
DB	4471	AAGGAAGCGAATTTTTCCTTCGGTCTAGCAGCCCGATTAAGTCCGAGTGCCTATTTT	4530
QY	3698	GTGCCAAAACGCTCATCCCACTGGCTCAAGGGTTTGAAGGGTTTCAATCGCCAAC	3757
DB	4531	GTGCCAAAACGCTCATCCCACTGGCTCAAGGGTTTGAAGGGTTTCAATCGCCAAC	4590
QY	3758	GATCGCCGACGTTTTCGCGACGTTTATTAATCTATATTTAAGTAGCTTTATTTGTT	3817
DB	4591	GATCGCCGACGTTTTCGCGACGTTTATTAATCTATATTTAAGTAGCTTTATTTGTT	4650
QY	3818	GTGTTTATGATTACAAAGTGATACACTAATTTTATAAAATTTATTTGAGTGGATTTT	3877
DB	4651	GTGTTTATGATTACAAAGTGATACACTAATTTTATAAAATTTATTTGAGTGGATTTT	4710
QY	3878	ATGTTGATTTCAGTAATCGAAAATAAGAGTTATGATTTCTCTGACAAAAGAGCAGATAA	3937
DB	4711	ATGTTGATTTCAGTAATCGAAAATAAGAGTTATGATTTCTCTGACAAAAGAGCAGATAA	4770
QY	3938	AAATTTAACAGATATCGCGCAACAAAAGGTTTTCAAAATCTCGGTTGCGGCTAGC	3997
DB	4771	AAATTTAACAGATATCGCGCAACAAAAGGTTTTCAAAATCTCGGTTGCGGCTAGC	4830
QY	3998	TATAGAAGATATGCAAGAAAGGAATCAGAACAAAATAAAGCGAAGCTCGCGTTT	4057
DB	4831	TATAGAAGATATGCAAGAAAGGAATCAGAACAAAATAAAGCGAAGCTCGCGTTT	4890
QY	4058	TAGAAGATATGCAAGTTTTCGCTACTGTTTGTGATGAAGTAA - TATATCATGCGCTATTAA	4116
DB	4891	TAGAAGATATGCAAGTTTTCGCTACTGTTTGTGATGAAGTAAATATATCATGCGCTATTAA	4950

RESULT 9

AAN91856/c

ID AAN91856 standard; DNA; 7336 BP.

XX

AAN91856;

AC

XX

25-MAR-2003 (revised)

31-OCT-2002 (revised)

23-MAR-1990 (first entry)

XX

DE

Plasmid pBHA-1.

6791 CAAACAGCTCTTCTACGATAGGCGACAAATCGCATCGTGGACGTTTGGCGTTCTACCG 6850
 2438 ATTTAGAGGTTTGATACACTTCTCTAGTATCCACCTGATCATATAATCGGCAATAG 2497
 6851 ATTTAGAGGTTTGATACACTTCTCTAGTATCCACCTGATCATATAATCGGCAATAG 6910
 2498 AGAAAAATTTGACCATGTGTAGCGGCGCAATCTGATTCACCTGAGATGCAATAATCTAGTA 2557
 6911 AGAAAAATTTGACCATGTGTAGCGGCGCAATCTGATTCACCTGAGATGCAATAATCTAGTA 6970
 2558 GAATCTCTCGCTATCAAAATTCAGTTCAGCTTCCAGTCCAGCTGCTCCATTCAGGC 2617
 6971 GAATCTCTCGCTATCAAAATTCAGTTCAGCTTCCAGTCCAGTCCAGTTCAGGC 7030
 2618 TGAATCTCTGCTCTCTGTTGACATGACACATCATCTCAATTCGAATAGGCGCCAT 2677
 7031 TGAATCTCTGCTCTCTGTTGACATGACACATCATCTCAATTCGAATAGGCGCCAT 7090
 2678 CAGTCTGACGACCAAGAGCGCATAAACACCAATAGCCTTAACATCATCCCATATTAT 2737
 7091 CAGTCTGACGACCAAGAGCGCATAAACACCAATAGCCTTAACATCATCCCATATTAT 7150
 2738 CCAATATTCTGCTCTTAAATTCATGAAACAACTTCTCTCTAGTCAATTATTA 2797
 7151 CCAATATTCTGCTCTTAAATTCATGAAACAACTTCTCTCTAGTCAATTATTA 7210
 2798 TTGGTCCATTCACTATTCTCATCTCCCTTTTCCAGATAATTTAGATTTGCTTTCTAAATA 2857
 7211 TTGGTCCATTCACTATTCTCATCTCCCTTTTCCAGATAATTTAGATTTGCTTTCTAAATA 7270
 2858 AGAATATTGGAGACCGCTTCTTATTCAGCTATTAATCTCTCTCTAGTCAATTATTA 2917
 7271 AGAATATTGGAGACCGCTTCTTATTCAGCTATTAATCTCTCTCTAGTCAATTATTA 7330
 2918 TTCAATCTCTTTTAAATAAATTTATAGCATCTAATCTTCAACAACTGCGCGTTTGTGA 2977
 7331 TTCAATCTCTTTTAAATAAATTTATAGCATCTAATCTTCAACAACTGCGCGTTTGTGA 7390
 2978 ACTACTCTTTTAAATAAATTTTCCGTTCCCAATTCACATTCGCAATATAGAAATC 3037
 7391 ACTACTCTTTTAAATAAATTTTCCGTTCCCAATTCACATTCGCAATATAGAAATC 7450
 3038 CATCTTCATCGGCTTTTTCGTCATCATCTGATGATCAATCAAACTGCGCTTCTCTGTGCAT 3097
 7451 CATCTTCATCGGCTTTTTCGTCATCATCTGATGATCAATCAAACTGCGCTTCTCTGTGCAT 7510
 3098 CAAGGTTTAAATTTTATGATTTCTTTTAAACCAACCAATAGAGATTAACTTTTAC 3157
 7511 CAAGGTTTAAATTTTATGATTTCTTTTAAACCAACCAATAGAGATTAACTTTTAC 7570
 3158 GGTGTAACCTTCTCCCAATCAGACAAACGTTTCAAAATCTTTTCTTCATCATCGGTCA 3217
 7571 GGTGTAACCTTCTCCCAATCAGACAAACGTTTCAAAATCTTTTCTTCATCATCGGTCA 7630
 3218 TAAATCCGATTCCTTTTACAGGATTTTTCAGTTTTCGTAATTCGCGATTTGATATCCG 3277
 7631 TAAATCCGATTCCTTTTACAGGATTTTTCAGTTTTCGTAATTCGCGATTTGATATCCG 7690
 3278 ATTATATTATTTTTCGTTATTTTATTAAGAGTCTCA 3318
 7691 ATTATATTATTTTTCGTTATTTTATTAAGAGTCTCA 7731

RESULT 11

AAT18951
 ID AAT18951 standard; DNA; 9144 BP.
 XX
 AC AAT18951;
 XX
 DT 17-JAN-1997 (first entry)
 XX
 DE Plasmid pBE346.
 XX

KW Spider; dragline protein; variant; monomer; polymer; circular;
 KW fibre forming region; Spidroin 1; Nephila clavipes; DPI; mimic;
 KW DP-1A analogue; fibre; high tensile strength; elasticity; clothing; rope;
 KW surgical suture; implant; reinforcement; film; coating; ss.
 XX
 OS Synthetic.
 PN MO9429450-A2.
 XX
 PD 22-DEC-1994.
 XX
 PF 15-JUN-1994; 94WO-US006689.
 XX
 PR 15-JUN-1993; 93US-00077600.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 PI Fahnstock SR;
 XX
 DR WPI; 1995-036479/05.
 XX
 PT New synthetic variants of spider dragline protein - for making fibres
 useful as clothing, surgical silk, plastic reinforcement etc., also
 related DNA, vectors and transformed cells.
 PT
 XX
 PS Example 6; Fig 14; 168pp; English.
 XX

This sequence represents the complete nucleotide sequence of the plasmid pBE346. This plasmid was used in the construction of the vector pP191 which was used to express synthetic spider dragline variants, DP-1 in B. subtilis. pBE346 comprises a replication origins that confer autonomous replication in both E. coli and B. subtilis, as well as antibiotic resistance makers selectable in E. coli (ampicillin) and B. subtilis (kanamycin). The plasmid also contains the lvs promoter and secretion signal operably linked to a staphylococcal protein A gene. The protein A gene is flanked by an EcoRV site at its proximal end, separating it from the lvs signal sequence, and a BamHI site at its distal end. The protein A gene is replaced by the DP-1 gene. The final plasmid pP191 was used to transform competent cells of B. subtilis B3010 (trp lys apr apr sacB) Kanamycin resistant transformants of DP-1 is a variant based the fibre forming regions of spider dragline protein, esp. the natural protein 1 (spidroin 1) from Nephila clavipes. The proteins may be used to produce fibres of high tensile strength and elasticity, suitable for clothing, rope, surgical sutures, biomaterials for implants, plastic reinforcements, films, coatings, etc

Sequence 9144 BP; 2656 A; 2150 C; 1790 G; 2548 T; 0 U; 0 Other;

Query Match 29.0%; Score 1493; DB 2; Length 9144;
 Best Local Similarity 98.1%; Pred. No. 3.2e-259;
 Matches 1511; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1778 TTCTTATCTTGATAATAAGGTAAGTAACTATTCGCGGATAGACTGTACATCTTCACGCATA 1837
 DB 6526 TACTCTTTTATATATCCCGACTGGCAATGCGCGGATAGACTGTACATCTTCACGCATA 6585
 QY 1838 AAATCCCTTTTCATTTTCTAATGTAATCTATTTACCTTATTATTATTAATTCCTCAT 1897
 DB 6586 AAATCCCTTTTCATTTTCTAATGTAATCTATTTACCTTATTATTATTAATTCCTCAT 6645
 QY 1898 AATTAATCCCTTTTCTTATTACGAAATGGCCGATTTAAGCACACCTTTATTCGGTT 1957
 DB 6646 AATTAATCCCTTTTCTTATTACGAAATGGCCGATTTAAGCACACCTTTATTCGGTT 6705
 QY 1958 AATGCGCCATGACGACCATGATAATTAATACTAGGAGAGTAAATAAATACGTAACC 2017
 DB 6706 AATGCGCCATGACGACCATGATAATTAATACTAGGAGAGTAAATAAATACGTAACC 6765
 QY 2018 AACATGATTAAATAATTTAGAGGTCATCGTTCAAAATGGTATGCGTTTTCACATCCA 2077
 DB 6766 AACATGATTAAATAATTTAGAGGTCATCGTTCAAAATGGTATGCGTTTTCACATCCA 6825

QY 2078 CTATATCGGTGCTGTTCTGTCCTCACTCTCGAATCCCAATCCAGAAATCTCTAGCGATT 2137
Db 6826 CTATATATCGGTGCTGTTCTGTCCTCACTCTCGAATCCCAATCCAGAAATCTCTAGCGATT 6885
QY 2138 CCAGAAATTTCTCAGAGTCGGAAGTTGACAGACATTTAGCAATCGCACAGATGTCAT 2197
Db 6886 CCAGAAATTTCTCAGAGTCGGAAGTTGACAGACATTTAGCAATCGCACAGATGTCAT 6945
QY 2198 AACCTGAAGGAAGATCTGATTGCTTAACTGCTTCACTGTTAGACCGGCTCGTCTGAT 2257
Db 6946 AACCTGAAGGAAGATCTGATTGCTTAACTGCTTCACTGTTAGACCGGCTCGTCTGAT 7005
QY 2258 AACAGATGCGATGATGAGACCAATCAATCAATGACCGCTGCTGATGCTGATGCTGATGCTA 2317
Db 7006 AACAGATGCGATGATGAGACCAATCAATCAATGACCGCTGCTGATGCTGATGCTGATGCTA 7065
QY 2318 AGGATGTAGAAATGTTGTCGGCTCTGACAGCAATTAATGACCAATTTGCGCTGCAATTT 2377
Db 7066 AGGATGTAGAAATGTTGTCGGCTCTGACAGCAATTAATGACCAATTTGCGCTGCAATTT 7125
QY 2378 CAAAACAGCTCTTCTACGATTAAGGCGCACAAATCGCATCTGGAAGTTTGGGCTTCTACCG 2437
Db 7126 CAAAACAGCTCTTCTACGATTAAGGCGCACAAATCGCATCTGGAAGTTTGGGCTTCTACCG 7185
QY 2438 ATTTAGCAGTTTGATACACTTCTCTAAGTATCCACCTGATCATTAATCGCAATATAG 2497
Db 7186 ATTTAGCAGTTTGATACACTTCTCTAAGTATCCACCTGATCATTAATCGCAATATAG 7245
QY 2498 AGAAAATTTGACCATGTTGTAAGCGGCGCAATCTGATCTGACCTGAGATGCAATCTAGTA 2557
Db 7246 AGAAAATTTGACCATGTTGTAAGCGGCGCAATCTGATCTGACCTGAGATGCAATCTAGTA 7305
QY 2558 GAATCTCTTCGCTATCAAAATTCATCTCCACCTTCCATCTCAATCGGTTGTCATTCATGCG 2617
Db 7306 GAATCTCTTCGCTATCAAAATTCATCTCCACCTTCCATCTCAATCGGTTGTCATTCATGCG 7365
QY 2618 TGAATCTGCTCTCTCTGACATGACACATCATCTCAATTCGAATAGGCGCCAT 2677
Db 7366 TGAATCTGCTCTCTCTGACATGACACATCATCTCAATTCGAATAGGCGCCAT 7425
QY 2678 CAGTCTGACGACCAAGAGAGCATAAACAACAATAGCTTAACTATCAATCCATTCATTTAT 2737
Db 7426 CAGTCTGACGACCAAGAGAGCATAAACAACAATAGCTTAACTATCAATCCATTCATTTAT 7485
QY 2738 CCAATATTCGTTCTTAATTTTCATGACAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2797
Db 7486 CCAATATTCGTTCTTAATTTTCATGACAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7545
QY 2798 TTGGTCCATTCATATCTCATTCCT 2857
Db 7546 TTGGTCCATTCATATCTCATTCCT 7605
QY 2858 AGAATATTTGGAGACACCGTTCTTATTCAGTATTAATTAATCTGCTTCTCTAGCATCC 2917
Db 7606 AGAATATTTGGAGACACCGTTCTTATTCAGTATTAATTAATCTGCTTCTCTAGCATCC 7665
QY 2918 TTCAATCTCTTTTAAATAAGATTTATAGCATCTCAATCTTCAAACTGCGCGTTGTTGA 2977
Db 7666 TTCAATCTCTTTTAAATAAGATTTATAGCATCTCAATCTTCAAACTGCGCGTTGTTGA 7725
QY 2978 ACTACTCTTTTAAATAAATTTTTCGGTTCCTATTCCTCAATTCATTCATTAATAGAAATC 3037
Db 7726 ACTACTCTTTTAAATAAATTTTTCGGTTCCTATTCCTCAATTCATTCATTAATAGAAATC 7785
QY 3038 CATCTTCATCGGCTTTTCGTCATCATCTGATGATCAATTCGCTTCTCTGTCAT 3097
Db 7786 CATCTTCATCGGCTTTTCGTCATCATCTGATGATCAATTCGCTTCTCTGTCAT 7845
QY 3098 CAAGGTTTAAATTTTATGATTTCTTTTAAACAAACCCACCATAGAGATTAACCTTTTAC 3157
Db 7846 CAAGGTTTAAATTTTATGATTTCTTTTAAACAAACCCACCATAGAGATTAACCTTTTAC 7905
QY 3158 GGTGTAAACCTTCTCTCAAAATCAGACAAAGTTTCAAAATCTTCTTCTCATCATCGGTCA 3217

Db 7906 GGTGTAAACCTTCTCTCAAAATCAGACAAAGTTTCAAAATCTTCTTCTCATCATCGGTCA 7965
QY 3218 TAAATCCGTATCTTTTACAGATATTTTGAGTTTCTGAGTTTCTCAATTCGCGATTGATATCCG 3277
Db 7966 TAAATCCGTATCTTTTACAGATATTTTGAGTTTCTGAGTTTCTCAATTCGCGATTGATATCCG 8025
QY 3278 ATTTATATTTATTTTCTGATTTTCTGATTTTCTGATTTTCTGATTTTCTGATTTTCTGATTT 3318
Db 8026 ATTTATATTTATTTTCTGATTTTCTGATTTTCTGATTTTCTGATTTTCTGATTTTCTGATTTT 8066

RESULT 12
AAT32232
ID AAT32232 standard; DNA; 10140 BP.
XX
AC AAT32232;
XX
DT 27-OCT-1996 (first entry)
XX
DE Plasmid pBE92.
XX
KW Thermostable enzyme; xylanase; xynA gene; beta-glucosidase; Bacillus;
KW Caldocellum saccharolyticum; polymerase chain reaction; PCR; primer;
KW signal peptide; pBE92; ss.
XX
OS Bacillus; sp.
OS Synthetic.
OS Chimeric.
XX
PN WO9623887-A1.
XX
PD 08-AUG-1996.
XX
PF 24-JAN-1996; 96WO-US000891.
XX
PR 30-JAN-1995; 95US-00380521.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Jackson EN, Liu G;
XX
DR WPI; 1996-371435/37.
XX
PT Thermostable enzymes e.g. xylanase and beta-glucosidase - are produced
PT extracellularly, in absence of generally required signal peptide, and in
PT high yields from recombinant bacteria.
XX
PS Example 1; Page 32-38; 48pp; English.
XX
CC Plasmid pBE92 (AAT32232) contains the Bacillus alkaline protease gene
CC promoter (aprp) and signal sequence (aprs). A PCR product (see also
CC AAT32230-31) contg. the Caldocellum saccharolyticum xylanase gene (xynA)
CC was ligated to the large fragment of pBE92. The resulting plasmid,
CC pBE158, contained the aprp-xynA fusion. Bacillus subtilis
CC transformants produced xylanase at 37 U/ml supernatant; this compared
CC with 272 and 1110 U/ml using constructs aprp-xynA (pBE145) (see also
CC AAT32225-26) and nprp-xynA (pBE145) (see also AAT32227), respectively,
CC which lacked a signal sequence
XX
SQ Sequence 10140 BP; 2821 A; 2414 C; 2129 G; 2776 T; 0 U; 0 Other;
Query Match 29.0%; Score 1493; DB 2; Length 10140;
Best Local Similarity 98.1%; Pred. No. 3.2e-259;
Matches 1511; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 1778 TTCTTATCTTGAATAAGGGTAACATTATTCGCGGATAGACTGTAACATTTCTCACGCATA 1837
Db 7523 TACTCTTTTAAATATCCCGACTGGCAATGCCGGATAGACGTACATTTCTCACGCATA 7582
QY 1838 AAATCCCTTCTATTTTCTAAATGTAATCTATTACCTTATTAATTAATTCGCTCAT 1897
Db 7583 AAATCCCTTCTATTTTCTAAATGTAATCTATTACCTTATTAATTAATTCGCTCAT 7642

XX	Eskelund MB, Schuelein M, Nielsen VS, Smets J;	
PI	WPI; 2003-059005/05.	
XX	New polypeptide useful as ingredient of detergent composition for	
PT	cleaning fabric, dishware or hard surface, encoded by DNA sequence	
PT	endogenous to strain of <i>Bacillus subtilis</i> , comprises pectate lyase	
PT	activity.	
XX	Disclosure; Page 84-87; 95pp; English.	
PS	The present invention relates to pectate lyase (EC 4.2.2.2) proteins and	
XX	polynucleotides encoding such proteins. The invention also relates to	
CC	detergent compositions comprising a surfactant and pectate lyase enzyme.	
CC	The detergent is useful for cleaning a fabric, a dishware or hard surface	
CC	to provide superior cleaning performance, for fabric cleaning and/or	
CC	fabric stain removal and/or fabric whiteness maintenance and/or fabric	
CC	softening and/or fabric colour appearance and/or fabric dye transfer	
CC	inhibition, for cleaning hard surfaces such as floors, walls or bathroom	
CC	tiles, for hand and machine dishwashing and for oral and/or dental	
CC	applications. They are useful for colour clarification, de-pilling and in	
CC	clay stain removal. The present sequence is pMOL995 plasmid DNA	
XX	Sequence 6661 BP; 1859 A; 1534 C; 991 G; 2277 T; 0 U; 0 Other;	
SQ	Query Match 29.0%; Score 1490.6; DB 7; Length 6661;	
	Best Local Similarity 99.1%; Pred. No. 8.2e-259;	
	Matches 1499; Conservative 0; Mismatches 14; Indels 0; Gaps 0;	
QY	1806 TGC CGG GAT A G C T G T A C A T T C T C A G C A T A A A A T C C C T T C A T T T C T A A T G T A A T	1865
DB	4072 T C C C G G G A T A G A C T G T A A C A T T C T C A C G C A T A A A T C C C C T T C A T T T C T A A T G T A A T	4131
QY	1866 C T A T T A C C T T A T T A T T A A T T C A A T T C G C T C A T A T T A A T C C T T T T C T T A T T A C G C A A A A	1925
DB	4132 C T A T T A C C T T A T T A T T A A T T C A A T T C G C T C A T A T T A A T C C T T T T C T T A T T A C G C A A A A	4191
QY	1926 T G C C C G G A T T A A G C A C A C C C T T A T T C C G T T A T T G C C C A T G A C A G C C A T G A A T T A C	1985
DB	4192 T G C C C G G A T T A A G C A C A C C C T T A T T C C G T T A T T G C C C A T G A C A G C C A T G A A T T A C	4251
QY	1986 T A A T A C T A G G A A G C T T A A T A A T A C G T A A C C A A C A T G A T T A A C A A T T A T T A G A G G T C A T	2045
DB	4252 T A A T A C T A G G A A G C T T A A T A A T A C G T A A C C A A C A T G A T T A A C A A T T A T T A G A G G T C A T	4311
QY	2046 C G T T C A A A T G G T A T G C G T T T T G A C A T P C C A T A T A T A T C C G T G C T T C T G C C A C T C	2105
DB	4312 C G T T C A A A T G G T A T G C G T T T T G A C A C A T C C A C T A T A T A T C C G T G C T T C T G T C C A C T C	4371
QY	2106 C T G A T C C C A T T C C A G A A T T C T C T A G C G A T T C C A G A A G T T T C T C A G A G T C G A A A G T T G	2165
DB	4372 C T G A T C C C A T T C C A G A A T T C T C T A G C G A T T C C A G A A G T T T C T C A G A G T C G A A A G T T G	4431
QY	2166 A C C A G A C A T T A C G A A C T G G C A C A G A T G G T C A T A P A C C T G A A G A A G A T C T G A T T G C T T A A C	2225
DB	4432 A C C A G A C A T T A C G A A C T G G C A C A G A T G G T C A T A P A C C T G A A G A A G A T C T G A T T G C T T A A C	4491
QY	2226 T G C T T C A G T T A G A C G A A G C C C T G T G T A T A A C A G A T G C G A T G C A T G C A G C A A A T C A A	2285
DB	4492 T G C T T C A G T T A G A C G A A G C C C T G T G T A T A A C A G A T G C G A T G C A T G C A G A C C A A T C A A	4551
QY	2286 C A T G G C A C T G C A T T G C T A C C T G T A C A G T C A A G G A T G T A G A A A T G T T G C G T C C T T G	2345
DB	4552 C A T G G C A C T G C A T T G C T A C C T G T A C A G T C A A G G A T G T A G A A A T G T T G C G T C C T T G	4611
QY	2346 C A C A G A A T A T T A C G C A T T T G C T G C A T A T T C A A C A G C T C T T C T A G A T A A G G G C A C A	2405
DB	4612 C A C A G A A T A T T A C G C A T T T G C T G C A T A T T C A A A C A G C T C T T C T A C G A T A A G G G C A C A	4671
QY	2406 A A T G C A T C G T G G A A G C T T T G G C C T T C A C C G A T T T A G C A G T T T G A T A C A C T T T C T C T A A	2465
DB	4672 A A T G C A T C G T G G A A G C T T T G G C C T T C T A C C G A T T T A G C A G T T T G A T A C A C T T T C T C T A A	4731

RESULT 14
AAA98012
ID AAA98012 standard; DNA; 5302 BP.
XX
XX AAA98012;
AC
XX
DT 05-FEB-2001 (first entry)
DE Expression vector p602/-CAT DNA.
XX
XX Lumazine synthase; capsid; cytostatic; antiviral; antibacterial; vaccine;

QY 2465 GTATCCACCTGAATCATATAATCGCAAAATAGAGAAAAATGACCATGTGTAAAGCGGCA 2525
DB 4732 GTATCCACCTGAATCATATAATCGCAAAATAGAGAAAAATGACCATGTGTAAAGCGGCA 4791
QY 2526 ATCTGATTCACCTGAGATGATATCTAGTAGAATCTCTCGCTATCAAAATTCACATTC 2585
DB 4792 ATCTGATTCACCTGAGATGATATCTAGTAGAATCTCTCGCTATCAAAATTCACATTC 4851
QY 2586 CACCTTCCACTCACCGGTTGTCCATTTCATGGCTGAACCTCTGCTCTCTGTGTGAATGAC 2645
DB 4852 CACCTTCCACTCACCGGTTGTCCATTTCATGGCTGAACCTCTGCTCTGTGTGAATGAC 4911
QY 2646 ACACATCATCTCAATATCGAATAGGCGCCATCTGTGACGACCAAGAGGCCATAAAC 2705
DB 4912 ACACATCATCTCAATATCGAATAGGCGCCATCTGTGACGACCAAGAGGCCATAAAC 4971
QY 2706 ACCAATAGGCTTAAACATCATCCCATATTTATCCCAATATTCCTTAAATTCATGAAC 2765
DB 4972 ACCAATAGGCTTAAACATCATCCCATATTTATCCCAATATTCCTTAAATTCATGAAC 5031
QY 2766 AATCTTCATCTCTCTCTAGTCAATATTTATTTGGTCAATTCACATATCTCATTCCTTT 2825
DB 5032 AATCTTCATCTCTCTCTAGTCAATATTTATTTGGTCAATTCACATATCTCATTCCTTT 5091
QY 2826 TTCAGATATTTTAGATTTTGTCTTTCTAATAAGAAATATTTGAGAGACACCGTCTCTATT 2885
DB 5092 TTCAGATATTTTAGATTTTGTCTTTCTAATAAGAAATATTTGAGAGACACCGTCTCTATT 5151
QY 2886 CAGCTATTAAATACCTGCTTCTCTAAGCATCTTCAATCCTTTAATAACAATATATAGCA 2945
DB 5152 CAGCTATTAAATACCTGCTTCTCTAAGCATCTTCAATCCTTTAATAACAATATATAGCA 5211
QY 2946 TCTAATCTTCAACAACATGCGCGTTTGTGAACCTCTTTAATAAAAAATATTTTCCG 3005
DB 5212 TCTAATCTTCAACAACATGCGCGTTTGTGAACCTCTTTAATAAAAAATATTTTCCG 5271
QY 3006 TTCCCAATCCCAATTCGAATAATAGAAAATCCATCTTCATCGGCTTTTTCGTCATCATC 3065
DB 5272 TTCCCAATCCCAATTCGAATAATAGAAAATCCATCTTCATCGGCTTTTTCGTCATCATC 5331
QY 3066 TGATGAATCAATCGCTCTCTCTGTGTCAATCAAGTTTAATTTTATGTATTTCTTT 3125
DB 5332 TGATGAATCAATCGCTCTCTCTGTGTCAATCAAGTTTAATTTTATGTATTTCTTT 5391
QY 3126 TAACAACACCATAGAGATTAACCTTTACGGTGTAAACCTTCCTCCAAATCAGACAA 3185
DB 5392 TAACAACACCATAGAGATTAACCTTTACGGTGTAAACCTTCCTCCAAATCAGACAA 5451
QY 3186 ACCTTTCAAAATCTTTTCTTCATCATCGGTCAATAAAATCCGTATCCTTTACAGGATATT 3245
DB 5452 ACCTTTCAAAATCTTTTCTTCATCATCGGTCAATAAAATCCGTATCCTTTACAGGATATT 5511
QY 3246 TGCAGTTTCGTCAATTCGGATTTGATATCCGATTTATATTTTATTTTTCGGTATTTTAA 3305
DB 5512 TGCAGTTTCGTCAATTCGGATTTGATATCCGATTTATATTTATTTTTCGGTCAATCAT 5571
QY 3306 TTAAAACGTCCTCA 3318
DB 5572 TTGAACTTTTACA 5584

XX gene therapy; immunotherapy; biosensor; diagnosis; ds.

XX Synthetic.

XX WO200053229-A2.

XX PN 14-SEP-2000.

XX PD 03-MAR-2000; 2000WO-EP001899.

XX PF 08-MAR-1999; 99DE-01010102.

XX PR (FISC/) FISCHER M.

XX PA (BACH/) BACHER A.

XX PI Fischer M, Bacher A;

XX XX WPI; 2000-572230/53.

XX Protein conjugate based on lumazine synthase as carrier, useful e.g. for
PT vaccination and immunotherapy, contains many functional molecules
PT attached to outside of the carrier.

XX Example 1; Page 130-132; 180pp; German.

CC This invention describes a novel protein conjugate (I) comprising at
CC least 1 functional region (FR) at any position in the sequence of a
CC carrier protein (II) to form a capsid three-dimensional structure of the
CC lumazine synthase (LS) type, such that the outer periphery is covalently
CC linked to many FR. The invention also describes (1) a hetero-oligomeric
CC protein conjugate (Ia) comprising either a mixture of at least 2
CC different (I) or of at least 1 (I) and at least 1 (II) that lacks FR.
CC with the components optionally covalently linked by chemical treatment;
CC (2) a method for preparation of (I) and (Ia); (3) a vector for
CC preparation of (I); (4) a DNA (III) that encodes (I); (5) LS from
CC Bacillus subtilis with Cys 93 and/or Cys139 replaced by Ser; (6) DNA
CC (IIa) encoding LS from Aquifex aeolicus which is codon-optimized for
CC expression in a recombinant Escherichia coli strain; (7) a chimeric
CC protein (CP) comprising amino acids (aa) 1-60 from LS of B. subtilis and
CC aa 61-154 from LS of A. aeolicus for use as (II); and (8) pharmaceutical
CC compositions and vaccines containing (I) and (Ia). The products of the
CC invention have cytostatic, antiviral and antibacterial activity and can
CC be used in a vaccine or for gene therapy. (I) are used as pharmaceuticals
CC (e.g. for immunotherapy of human immune deficiency virus infection or
CC tumors) and in vaccines, including multivalent vaccines, against
CC bacterial or viral infections to produce diagnostic or therapeutic
CC antibodies, for selective detection, purification and characterization of
CC antibodies, and for preparation of protein libraries. (I) may also be
CC used as biosensors and for diagnosis, e.g. of tumors. Genes that encode
CC (I) are useful in DNA vaccines and for preparation of plant-based oral
CC vaccines. (I) can contain many FR (same or different) at the surface of a
CC spherical particle (LS comprises 60 subunits that assemble into an
CC icosahedron). The large number of FR may increase sensitivity in
CC immunoassays and the efficiency of immunotherapy agents

XX Sequence 5302 BP; 1502 A; 1208 C; 962 G; 1630 T; 0 U; 0 Other;

XX Query Match 28.9%; Score 1487.6; DB 3; Length 5302;

XX Best Local Similarity 99.1%; Pred. No. 2.8e-258;

XX Matches 1496; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1809 CGGATAGACTGTAACTCTCAGCATATAAATCCCTTTTCATTTCTAATGTAAATCTTA 1868

Db 2716 CGGATAGACTGTAACTCTCAGCATATAAATCCCTTTTCATTTCTAATGTAAATCTTA 2775

QY 1869 TTACCTTTAATTAATTCATTCGCTCATATAATATCTTTTCTTATTACGCAAAATGG 1928

Db 2776 TTACCTTTAATTAATTCATTCGCTCATATAATATCTTTTCTTATTACGCAAAATGG 2835

QY 1929 CCCGATTTAAGACACACCTTTTATTCGGTTAATGGCCATGACGCCATGATTAATCTAA 1988

Db 2836 CCCGATTTAAGACACACCTTTTATTCGGTTAATGGCCATGACGCCATGATTAATCTAA 2895

QY 1989 TACTAGGAGAGCTTAATAATAACGTAAACCAACATGATTAAACAATTTATTAGAGGTCACTGT 2048

Db 2896 TACTAGGAGAGCTTAATAATAACGTAAACCAACATGATTAAACAATTTATTAGAGGTCACTGT 2955

QY 2049 TCAAAATGGTATGCGTTTGTACACATCCACATATATATCCGTTGCTGTCCTGCTCCACTCCTG 2108

Db 2956 TCAAAATGGTATGCGTTTGTACACATCCACATATATATCCGTTGCTGTCCTGCTCCACTCCTG 3015

QY 2109 AATCCCATTTCCAGAAATTTCTTAGCGATTCCAGAGTTTCTCAGAGTCGGAAGTTGACC 2168

Db 3016 AATCCCATTTCCAGAAATTTCTTAGCGATTCCAGAGTTTCTCAGAGTCGGAAGTTGACC 3075

QY 2169 AGACATTAGCACTGGCAGACAGATGCTCATACCTCAAGGAAGATCTGATTCTTAACCTGC 2228

Db 3076 AGACATTAGCACTGGCAGACAGATGCTCATACCTCAAGGAAGATCTGATTCTTAACCTGC 3135

QY 2229 TTTCAAGTAAAGACCGAAGCGCTCGTGTATTAACAGATGCGATGCGATGATGATGATGATGAT 2288

Db 3136 TTTCAAGTAAAGACCGAAGCGCTCGTGTATTAACAGATGCGATGCGATGATGATGATGATGAT 3195

QY 2289 GGCACCTGCAATGCTACCTGTACAGTCAAGATGCTAGAAATGTTGCGTTCCTTGAC 2348

Db 3196 GGCACCTGCAATGCTACCTGTACAGTCAAGATGCTAGAAATGTTGCGTTCCTTGAC 3255

QY 2349 ACGAATATTACGCCATTTGCGTGTACATATTCAACAGCTCTTCTACGATAAGGCGACAAT 2408

Db 3256 ACGAATATTACGCCATTTGCGTGTACATATTCAACAGCTCTTCTACGATAAGGCGACAAT 3315

QY 2409 CGCATCGTGAAGACGTTTGGGCTTCTACCGATTAGCAGTTTGATACACTTCTCTAAGTA 2468

Db 3316 CGCATCGTGAAGACGTTTGGGCTTCTACCGATTAGCAGTTTGATACACTTCTCTAAGTA 3375

QY 2469 TCCACTCGAATCAATAATCGCAATAGAGAAATTTGACCATGTGTAAAGCGCAATC 2528

Db 3376 TCCACTCGAATCAATAATCGCAATAGAGAAATTTGACCATGTGTAAAGCGCAATC 3435

QY 2529 TGATTCCACCTGAGATGATATCTAGTAGAATCTCTTCGCTATCAAAATTTCACTTCCAC 2588

Db 3436 TGATTCCACCTGAGATGATATCTAGTAGAATCTCTTCGCTATCAAAATTTCACTTCCAC 3495

QY 2589 CTTCACTACCGGTTGCTCATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2648

Db 3496 CTTCACTACCGGTTGCTCATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3555

QY 2649 CATCATCTCAATATCCGAATAGGCGCCATCAGTCTGACGACCAAGAGAGCCATTAACACC 2708

Db 3556 CATCATCTCAATATCCGAATAGGCGCCATCAGTCTGACGACCAAGAGAGCCATTAACACC 3615

QY 2709 AATAGCCTTAACATCATCCCATATTTATCAATATTCGTTTCTTATTTTATGAGCAAT 2768

Db 3616 AATAGCCTTAACATCATCCCATATTTATCAATATTCGTTTCTTATTTTATGAGCAAT 3675

QY 2769 CTTCACTCTTCTCTCTAGTCATTATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2828

Db 3676 CTTCACTCTTCTCTCTAGTCATTATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3735

QY 2829 AGATAATTTTAGATTTGCTTTCTTAATAAGAAATTTTGGAGAGCAGCGTTCTTATTTCAG 2888

Db 3736 AGATAATTTTAGATTTGCTTTCTTAATAAGAAATTTTGGAGAGCAGCGTTCTTATTTCAG 3795

QY 2889 CTATTAAATCACTGCTTCTTCCCTAAGCATCTTCAATCTCTTTTAAATAAATTTATGACATCT 2948

Db 3796 CTATTAAATCACTGCTTCTTCCCTAAGCATCTTCAATCTCTTTTAAATAAATTTATGACATCT 3855

QY 2949 AATCTTCAACAACTGGCCGCTTTTGTGCACTACTCTTTTAAATAAATTTTCCGTTTC 3008

Db 3856 AATCTTCAACAACTGGCCGCTTTTGTGCACTACTCTTTTAAATAAATTTTCCGTTTC 3915

QY 3009 CCAATTCACATTTGAATAATAGAAATCCCATCTTCATCGGCTTTTTCGTCATCATCTGT 3068

Db 3916 CCAATTCACATTTGAATAATAGAAATCCCATCTTCATCGGCTTTTTCGTCATCATCTGT 3975

QY 3069 ATGAATCAAAATCGCCTTCTTCTGTGTCATCAAGGTTTAAATTTTATGATTTCTTTTAA 3128

Db 3976 ATGAATCAATCGCGCTTCTCTGTGTCATCAAGGTTAAATTTTATGTATTTCTTTTAA 4035
Qy 3129 CAACCCACCATAGAGATTAACTTTTACGGTGTAAACCTTCCCAAATCAGACAAACG 3188
Db 4036 CAACCCACCATAGAGATTAACTTTTACGGTGTAAACCTTCCCAAATCAGACAAACG 4095
Qy 3189 TTTCAAATCTTTCTTTCATCATCGGTCAATAAAATCCGGTATCCCTTTACAGGATTTTGC 3248
Db 4096 TTTCAAATCTTTCTTTCATCATCGGTCAATAAAATCCGGTATCCCTTTACAGGATTTTGC 4155
Qy 3249 AGTTTCGTCATTCGCGATTGATATCCGATTATATTTTTCGTTATTTTATTA 3308
Db 4156 AGTTTCGTCATTCGCGATTGATATCCGATTATATTTTTCGTTTCGTCGATCAATTTG 4215
Qy 3309 AAACGTCCTCA 3318
Db 4216 AACTTTTACA 4225
RESULT 15
AAA98014
ID AAA98014 standard; DNA; 5767 BP.
AC AAA98014;
XX 05-FEB-2001 (first entry)
DT Expression vector p602-BS-LuSy DNA.
DE Lumazine synthase; capsid; cytostatic; antiviral; antibacterial; vaccine;
KW gene therapy; immunotherapy; biosensor; diagnosis; ds.
XX Synthetic.
XX WO200053229-A2.
XX 14-SEP-2000.
XX 03-MAR-2000; 2000WO-EP001899.
XX 08-MAR-1999; 99DE-01010102.
XX (FISC/) FISCHER M.
XX (BACH/) BACHER A.
XX Fischer M, Bacher A;
XX WPI; 2000-572230/53.
XX Protein conjugate based on lumazine synthase as carrier, useful e.g. for
XX vaccination and immunotherapy, contains many functional molecules
XX attached to outside of the carrier.
XX Example 2; Page 134-137; 180pp; German.
XX This invention describes a novel protein conjugate (I) comprising at
XX least 1 functional region (FR) at any position in the sequence of a
XX carrier protein (II) to form a capsid three-dimensional structure of the
XX lumazine synthase (LS) type, such that the outer periphery is covalently
XX linked to many FR. The invention also describes (1) a hetero-oligomeric
XX protein conjugate (Ia) comprising either a mixture of at least 2
XX different (i) or of at least 1 (ii) and at least 1 (iii) that lacks FR,
XX with the components optionally covalently linked by chemical treatment;
XX (2) a method for preparation of (I) and (Ia); (3) a vector for
XX preparation of (I); (4) a DNA (III) that encodes (I); (5) LS from
XX Bacillus subtilis with Cys 93 and/or Cys139 replaced by Ser; (6) DNA
XX (IIia) encoding LS from Aquifex aeolicus which is codon-optimized for
XX expression in a recombinant Escherichia coli strain; (7) a chimeric
XX protein (Cp) comprising amino acids (aa) 1-60 from LS of B. subtilis and
XX aa 61-154 from LS of A. aeolicus for use as (Iii); and (8) pharmaceutical
XX compositions and vaccines containing (I) and (Ia). The products of the
XX invention have cytostatic, antiviral and antibacterial activity and can

CC be used in a vaccine or for gene therapy. (I) are used as pharmaceuticals
CC (e.g. for immunotherapy of human immune deficiency virus infection or
CC tumors) and in vaccines, including multivalent vaccines, against
CC bacterial or viral infections, to produce diagnostic or therapeutic
CC antibodies, for selective detection, purification and characterization of
CC antibodies, and for preparation of protein libraries. (I) may also be
CC used as biosensors and for diagnosis, e.g. of tumors. Genes that encode
CC (I) are useful in DNA vaccines and for preparation of plant-based oral
CC vaccines. (I) can contain many FR (same or different) at the surface of a
CC spherical particle (US comprises 60 subunits that assemble into an
CC icosahedron). The large number of FR may increase sensitivity in
CC immunoassays and the efficiency of immunotherapy agents

XX Sequence 5767 BP; 1654 A; 1297 C; 1073 G; 1743 T; 0 U; 0 Other;

Query Match 28.9%; Score 1487.6; DB 3; Length 5767;
Best Local Similarity 99.1%; Pred. No. 2.8e-258;
Matches 1496; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1809 CGGATAGACTGTAAACATTTCTCAGCATATAAATCCCTTTTCATTTTCTTAATGTAATCTA 1868
Db 3181 CGGATAGACTGTAAACATTTCTCAGCATATAAATCCCTTTTCATTTTCTTAATGTAATCTA 3240
Qy 1869 TTACCTTATTATTAAATCAATTCGCTCATAATTAATCCCTTTTCTTATTACGCAAAATGG 1928
Db 3241 TTACCTTATTATTAAATCAATTCGCTCATAATTAATCCCTTTTCTTATTACGCAAAATGG 3300
Qy 1929 CCGATTAAAGCACACCCCTTATTCGGTTAAATCGCCATGACAGCCATGATTAATTAATA 1988
Db 3301 CCGATTAAAGCACACCCCTTATTCGGTTAAATCGCCATGACAGCCATGATTAATTAATA 3360
Qy 1989 TACTAGAGAAAGTTAAATAATACGTAAACCAACATGATTAACCAATTAATAGAGTCTACGT 2048
Db 3361 TACTAGAGAAAGTTAAATAATACGTAAACCAACATGATTAACCAATTAATAGAGTCTACGT 3420
Qy 2049 TCAAAATGATCGTTTTCAGACATCCACTATATATCGTCTCGTCTCTGTCACACTCCTG 2108
Db 3421 TCAAAATGATCGTTTTCAGACATCCACTATATATCGTCTCGTCTCTGTCACACTCCTG 3480
Qy 2109 AATCCCATCCAGAAATTTCTTAGCGATTCCAGAGTTTCTCAGAGTCGAAAGTTGACC 2168
Db 3481 AATCCCATCCAGAAATTTCTTAGCGATTCCAGAGTTTCTCAGAGTCGAAAGTTGACC 3540
Qy 2169 AGACATTACGAATCGGCACAGATGGTCAATAACCTGAAGGAAGATCTGATTGCTTAATGTC 2228
Db 3541 AGACATTACGAATCGGCACAGATGGTCAATAACCTGAAGGAAGATCTGATTGCTTAATGTC 3600
Qy 2229 TTCAGTTAAGACCGAAGCGCTCGTGTATAACAGATCGGATGATGAGACCAATCAACAT 2288
Db 3601 TTCAGTTAAGACCGAAGCGCTCGTGTATAACAGATCGGATGATGAGACCAATCAACAT 3660
Qy 2289 GGCACCTGCCATTTGCTACCTGTACAGTCAAGGATGGTAGAAATGTTGTGCGTCTTGCAC 2348
Db 3661 GGCACCTGCCATTTGCTACCTGTACAGTCAAGGATGGTAGAAATGTTGTGCGTCTTGCAC 3720
Qy 2349 ACGAATATTAGCCCATTTGCTGCTCATATTCAACACAGCTCTTCTACGATAGGCGCAAAAT 2408
Db 3721 ACGAATATTAGCCCATTTGCTGCTCATATTCAACACAGCTCTTCTACGATAGGCGCAAAAT 3780
Qy 2409 CGCATCGTGAAACGTTTGGGCTTCTACCGAGTTTACGAGTTTGTATACACATTTTCTTAAGTA 2468
Db 3781 CGCATCGTGAAACGTTTGGGCTTCTACCGAGTTTACGAGTTTGTATACACATTTTCTTAAGTA 3840
Qy 2469 TCCACCTGAATCAATAATCGGCAAAATAGAGAAAATTTAGCAATGTTGTAAGCGGCCAATC 2528
Db 3841 TCCACCTGAATCAATAATCGGCAAAATAGAGAAAATTTAGCAATGTTGTAAGCGGCCAATC 3900
Qy 2529 TGATTCACCTGAGATGCAATAATCTAGTAGAAATCTCTTCGCTATCAAAATTTCACTTCCAC 2588
Db 3901 TGATTCACCTGAGATGCAATAATCTAGTAGAAATCTCTTCGCTATCAAAATTTCACTTCCAC 3960
Qy 2589 CTTCACCTACCGGTTGTCATTCATGGCTGAATCTGCTTCTCTCTGTTGACATGACACA 2648

Db	3961	CTTCCACTCACGGTTGTCATTATGCGTGAACCTCTGCTTCCTCTGTTGACATGACACA	4020
Qy	2649	CATCATCTCAATATCCGATAGGCGCCATCAGTCTGACGACCAAGAGAGCCATAAACACC	2708
Db	4021	CATCATCTCAATATCCGATAGGCGCCATCAGTCTGACGACCAAGAGAGCCATAAACACC	4080
Qy	2709	AATAGCCTTAACATCATCCCATATTTATCCAAATATTCGTTCCCTTAATTTATGAACAAT	2768
Db	4081	AATAGCCTTAACATCATCCCATATTTATCCAAATATTCGTTCCCTTAATTTATGAACAAT	4140
Qy	2769	CTTCAATCTTTCTCTCTAGTCATATATTTATGTTCCATTCACATATCTCATTCCTTTTC	2828
Db	4141	CTTCAATCTTTCTCTCTAGTCATATATTTATGTTCCATTCACATATCTCATTCCTTTTC	4200
Qy	2829	AGATAATTTTAGATTTGCTTTTCTAAATAAGAAATATTTGAGAGACCGTTCTTATTACAG	2888
Db	4201	AGATAATTTTAGATTTGCTTTTCTAAATAAGAAATATTTGAGAGACCGTTCTTATTACAG	4260
Qy	2889	CTATTAAATCACTCGTCTTCTTAAGCATCCTTCAATCCTTTTAAATAACAAATATAGCATCT	2948
Db	4261	CTATTAAATCACTCGTCTTCTTAAGCATCCTTCAATCCTTTTAAATAACAAATATAGCATCT	4320
Qy	2949	RATCTTCAACAACTGGCGCGTTGTTGCACTACTCTTTTAAATAAAATTTTCCGTTTC	3008
Db	4321	RATCTTCAACAACTGGCGCGTTGTTGCACTACTCTTTTAAATAAAATTTTCCGTTTC	4380
Qy	3009	CCAAATCCACATTTGCAATATAGAAATCCATCTTCATPCGGCTTTTTCGTCATCATCTGT	3068
Db	4381	CCAAATCCACATTTGCAATATAGAAATCCATCTTCATPCGGCTTTTTCGTCATCATCTGT	4440
Qy	3069	ATGAATCAAAATCGCCTTCTCTGTCATCAAGGTTTAAATTTTATGTTATCTTTTAA	3128
Db	4441	ATGAATCAAAATCGCCTTCTCTGTCATCAAGGTTTAAATTTTATGTTATCTTTTAA	4500
Qy	3129	CAAAACCATAGGAGATTAACTTTTACGGTGTAACCTTCTCCCAAATCAGACAAACG	3188
Db	4501	CAAAACCATAGGAGATTAACTTTTACGGTGTAACCTTCTCCCAAATCAGACAAACG	4560
Qy	3189	TTTCAAAATCTTTCTTCTCATATCGGTCATAAAATCCGTAATCCTTTACAGGATTTTTC	3248
Db	4561	TTTCAAAATCTTTCTTCTCATATCGGTCATAAAATCCGTAATCCTTTACAGGATTTTTC	4620
Qy	3249	AGTTTCGTCAAATTCGGATTTGATATCCGATTTATATTTTTCGTTATTTTATTATTA	3308
Db	4621	AGTTTCGTCAAATTCGGATTTGATATCCGATTTATATTTTTCGTTATTTTATTATTA	4680
Qy	3309	AAACGTCCTCA	3318
Db	4681	AACTTTTACA	4690

Search completed: September 23, 2004, 19:07:26
Job time : 1178.56 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 23, 2004, 11:31:21 ; Search time 12724.2 Seconds
(without alignments)
17515.470 Million cell updates/sec

Title: US-10-030-390-1
Perfect score: 5142
Sequence: 1 gaattcgagctggtaccg.....ccgtctgtgcttcttaag 5142

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_atg.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_on.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vl.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	5142	100.0	5142	6	AX069289	AX069289 Sequence
C 2	2467.2	48.0	8119	6	AR002333	AR002333 Sequence
C 3	2467.2	48.0	8119	6	AR018032	AR018032 Sequence
C 4	2467.2	48.0	8119	6	AR055698	AR055698 Sequence
C 5	1650.4	32.1	2059	6	A09338	A09338 S.lactis pl
C 6	1650.4	32.1	2059	6	A09339	A09339 S.lactis pl
C 7	1641.6	31.9	8082	12	AF076212	AF076212 Promoter
C 8	1641.6	31.9	8142	12	AF076213	AF076213 Promoter
C 9	1637	31.8	2178	1	LR38A	X56954 L.lactis pl
C 10	1635.2	31.8	5126	12	AF372620	AF372620 GlnQ allele
C 11	1635.2	31.8	9021	12	AF061336	AF061336 Tn317 del
C 12	1635.2	31.8	13561	12	AY028776	AY028776 TnpA2 mu
C 13	1629.6	31.7	2505	1	LR38A	X54310 Lactococcus
C 14	1575.2	30.6	4438	12	AB055651	AB055651 Thermosen
C 15	1575.2	30.6	4438	12	AB055652	AB055652 Thermosen
C 16	1575.2	30.6	4506	12	AB055650	AB055650 Thermosen
C 17	1572.2	30.6	3522	12	AB084785	AB084785 Transpos
C 18	1572.2	30.6	3618	12	AB084784	AB084784 Transpos
C 19	1493	29.0	4545	1	PR0110CG	M37273 Plasmid PUB
C 20	1493	29.0	4548	1	PR0110CG	M19465 Plasmid PUB
C 21	1493	29.0	7336	6	A08030	A08030 Synthetic n
C 22	1493	29.0	7336	6	A13198	A13198 complete pl
C 23	1493	29.0	7336	6	A20502	A20502 Hind III fr
C 24	1493	29.0	7336	6	AR033984	AR033984 Sequence
C 25	1493	29.0	7336	6	AR123993	AR123993 Sequence
C 26	1493	29.0	7336	6	AR154323	AR154323 Sequence
C 27	1493	29.0	7336	6	AR408644	AR408644 Sequence
C 28	1491.4	29.0	4525	1	BSREP	X15670 Bacillus pl
C 29	1491.4	29.0	6447	12	AF288421	AF288421 Synthetic
C 30	1491.4	29.0	7326	6	AR364483	AR364483 Sequence
C 31	1491.4	29.0	11887	1	PTITERIMOB	M63891 Bacillus st
C 32	1491	29.0	7336	6	A05160	A05160 Nucleotide
C 33	1490.6	29.0	6661	6	AX601433	AX601433 Sequence
C 34	1490.2	29.0	4780	12	ACVPBD64	X51450 Artificial
C 35	1487.6	28.9	5767	6	AX035965	AX035965 Sequence
C 36	1482.6	28.8	2763	6	E05086	E05086 DBA encodin
C 37	1482.6	28.8	2763	6	E05087	E05087 Total seque
C 38	1479	28.8	5609	6	AR369757	AR369757 Sequence
C 39	1479	28.8	5609	6	AR369758	AR369758 Sequence
C 40	1479	28.8	7026	6	AR369761	AR369761 Sequence
C 41	1479	28.8	7344	6	AR369759	AR369759 Sequence
C 42	1436.2	27.9	1600	1	LPCOPAREP	X95843 L.plantarum
C 43	1386.6	27.0	5126	12	AF372620	AF372620 GlnQ allele
C 44	1336.6	26.0	2466	1	SA110KAR	X03408 Staphylococ
C 45	1335.6	26.0	2262	1	BSPRBHK	X03409 Thermophil

ALIGNMENTS

RESULT 1

AX069289

LOCUS

DEFINITION

AX069289

ACCESSION

VERSION

AX069289.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1

AUTHORS

TITLE

JOURNAL

AX069289 Sequence 1 from Patent WO0102570.
5142 bp DNA linear PAT 25-JAN-2001

AX069289.1 GI:12579162

Synthetic construct
synthetic construct
artificial sequences.

Hans.W.C., Steidler,L. and Remaut,E.R.
Delivery of trefoil peptides
Patent: WO 0102570-A 111-JAN-2001.
Vlaams Interuniversitair Instituut voor Biotechnologie (BE)

Pred. No. is the number of results predicted by chance to have a

FEATURES		Location/Qualifiers	
source		1..5142	
		/organism="synthetic construct" /mol_type="unassigned DNA" /db_xref="taxon:32630" /note="Synthetic construct : plasmid"	
ORIGIN			
Query Match	100.0%;	Score 5142;	DB 6; Length 5142;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 5142;	Conservative	0; Mismatches	0; Indels 0; Gaps
Qy	1	GAATTCGAGCTCGGTACCCGGGGATCTCGATCCCGCGAAATTAATACGACTCACTATAGG	60
Db	1	GAATTCGAGCTCGGTACCCGGGGATCTCGATCCCGCGAAATTAATACGACTCACTATAGG	60
Qy	61	GAGACCACAACGGTTTCCCTCTAGAAATAAATTTGTGTTAACTTTAAGAGGAGATATACA	120
Db	61	GAGACCACAACGGTTTCCCTCTAGAAATAAATTTGTGTTAACTTTAAGAGGAGATATACA	120
Qy	121	TATGAAAAAAGATATCTCAGCTATTTTAATGTCACAGTCATACTTTCCTGCTGCAGC	180
Db	121	TATGAAAAAAGATATCTCAGCTATTTTAATGTCACAGTCATACTTTCCTGCTGCAGC	180
Qy	181	CCCGTTGTCAGGTGTTTACGCCAGGCCAGGCCCGCCAGGAGAAACATGTATCAT	240
Db	181	CCCGTTGTCAGGTGTTTACGCCAGGCCAGGCCCGCCAGGAGAAACATGTATCAT	240
Qy	241	GGCCCCCGGGAGAGGATAAATTTGGCTTCCCGGTGTCAACGCCAGCAGTGCAACGA	300
Db	241	GGCCCCCGGGAGAGGATAAATTTGGCTTCCCGGTGTCAACGCCAGCAGTGCAACGA	300
Qy	301	GAGAGTTGCTGTTTTTGATGACAGTGTCCGGGGATTCCTCGTGTGCTTCCACCCCATGCG	360
Db	301	GAGAGTTGCTGTTTTTGATGACAGTGTCCGGGGATTCCTCGTGTGCTTCCACCCCATGCG	360
Qy	361	CATCGAGAACACTCAAGAAGAGAAATGCCCTTCTAACTAGTAGATCCGGCTGCTAACAA	420
Db	361	CATCGAGAACACTCAAGAAGAGAAATGCCCTTCTAACTAGTAGATCCGGCTGCTAACAA	420
Qy	421	AGCCCGAAGGAAGCTGAGTTGGCTGTGCAACCGCTGAGCATAACTAGCATAAACCCCT	480
Db	421	AGCCCGAAGGAAGCTGAGTTGGCTGTGCAACCGCTGAGCATAACTAGCATAAACCCCT	480
Qy	481	TGGGGCCTCTAAACGGGCTTTGAGGGGTTTTTTCCTCGAAAGGAGGAACATATATCCGGATG	540
Db	481	TGGGGCCTCTAAACGGGCTTTGAGGGGTTTTTTCCTCGAAAGGAGGAACATATATCCGGATG	540
Qy	541	ACCTGCAGCATGCAAGCTTGGCACTGGCGTCTGTTTTACAAAGTGGTGCATGATTTCAAC	600
Db	541	ACCTGCAGCATGCAAGCTTGGCACTGGCGTCTGTTTTACAAAGTGGTGCATGATTTCAAC	600
Qy	601	CCTGGGCTTACCCAACTTAATCGCCTTCGCAGCACATCCCTTTTCGCCAGCTCATTTCCAC	660
Db	601	CCTGGGCTTACCCAACTTAATCGCCTTCGCAGCACATCCCTTTTCGCCAGCTCATTTCCAC	660
Qy	661	TTTTTGCACTTACAAACTGCANAACTCATATGTAATAATTCGTCCTTTTAGTGGGCACAA	720
Db	661	TTTTTGCACTTACAAACTGCANAACTCATATGTAATAATTCGTCCTTTTAGTGGGCACAA	720
Qy	721	ATGTGAGGCATTTTCGCTCTTTCCGGCGAGCTAGTTACCTTAAAGTTATTTGGTATGACT	780
Db	721	ATGTGAGGCATTTTCGCTCTTTCCGGCGAGCTAGTTACCTTAAAGTTATTTGGTATGACT	780
Qy	781	GGTTTTAAGCGCAAAAAAAGTTGCTTTTCGTACCTATTAATGTATCGTTTTTAGAAAAAC	840
Db	781	GGTTTTAAGCGCAAAAAAAGTTGCTTTTCGTACCTATTAATGTATCGTTTTTAGAAAAAC	840
Qy	841	GACTGTAAAGGTACAGTCGGCATTCCTCATATTAATAAGCCAGTCATTAGGCCCTATC	900
Db	841	GACTGTAAAGGTACAGTCGGCATTCCTCATATTAATAAGCCAGTCATTAGGCCCTATC	900
Qy	901	TGCAATTCCTGAATAGAGTTTCATAAAACAATCTCGCATGATAACCATCAAAACGAATG	960

Db 1981 ATTACTAATACTAGGAGAGTTAATAAATACGTAAACCAACATGATTAAACAATTTAGAG 2040
Qy
2041 GTCAATCGTTCAAAATGGTATGCGTTTGTGACACATCCACTATATATCCGTGTCGTCTGTC 2100
Db
2041 GTCAATCGTTCAAAATGGTATGCGTTTGTGACACATCCACTATATATCCGTGTCGTCTGTC 2100
Qy
2101 CACTCTGGAATCCGATCCGAGAAATCTCTAGCGATCCAGAGTTTCTCAGAGTCGGA 2160
Db
2101 CACTCTGGAATCCGATCCGAGAAATCTCTAGCGATCCGAGAGTTTCTCAGAGTCGGA 2160
Qy
2161 AGTTGACGACATTTAGCACTGGCAGAGATGGTCATTAACCTGAAGGAGATCTGATTGC 2220
Db
2161 AGTTGACGACATTTAGCACTGGCAGAGATGGTCATTAACCTGAAGGAGATCTGATTGC 2220
Qy
2221 TTAACCTGCTTCAAGTAAAGACGAGCGCTCGTGTATTAACAGATGGATCCAGACCA 2280
Db
2221 TTAACCTGCTTCAAGTAAAGACGAGCGCTCGTGTATTAACAGATGGATCCAGACCA 2280
Qy
2281 ATCAACATGACCTGCTGCTTACCTGTACAGTCAAGGATGGTGAATGTTGTCGT 2340
Db
2281 ATCAACATGACCTGCTGCTTACCTGTACAGTCAAGGATGGTGAATGTTGTCGT 2340
Qy
2341 CTTTGCACAGCAATATTACGCCATTTGCCCTGCATATTCAAACAGCTCTTCTAGCATAAG 2400
Db
2341 CTTTGCACAGCAATATTACGCCATTTGCCCTGCATATTCAAACAGCTCTTCTAGCATAAG 2400
Qy
2401 GCACAAATCGCATCGTGGAGAGTTGGCTTCTACCGATTTAGCAGTTTGCATACCTTC 2460
Db
2401 GCACAAATCGCATCGTGGAGAGTTGGCTTCTACCGATTTAGCAGTTTGCATACCTTC 2460
Qy
2461 TCTAAGTATCCACTGGAATCATAAATCGGCAAAATAGAGAAATGACCATGTGAAGC 2520
Db
2461 TCTAAGTATCCACTGGAATCATAAATCGGCAAAATAGAGAAATGACCATGTGAAGC 2520
Qy
2521 GGCCAAATCTGATTCGACTGAGATGCAATCTAGTAGAATCTCTCGCTATCAAAATTC 2580
Db
2521 GGCCAAATCTGATTCGACTGAGATGCAATCTAGTAGAATCTCTCGCTATCAAAATTC 2580
Qy
2581 ACTTCCACCTTCCACTCAGCGTTGTCATTCAGGCTGAACCTGCTTCTCTGTTGAC 2640
Db
2581 ACTTCCACCTTCCACTCAGCGTTGTCATTCAGGCTGAACCTGCTTCTCTGTTGAC 2640
Qy
2641 ATGACACATPCATCTCAATATCGAATAGGCGCCATCAGTCTGACCAAGAGAGCCA 2700
Db
2641 ATGACACATPCATCTCAATATCGAATAGGCGCCATCAGTCTGACCAAGAGAGCCA 2700
Qy
2701 TAAACACCAATAGCTTAACATATCCCAATTTATCCCAATTCGTTCTCTTAATTTCA 2760
Db
2701 TAAACACCAATAGCTTAACATATCCCAATTTATCCCAATTCGTTCTCTTAATTTCA 2760
Qy
2761 TGAACAACTCTCAATCTTCTCTAGTCAATTTATTTGGTCCATTCACATTTCTCAT 2820
Db
2761 TGAACAACTCTCAATCTTCTCTAGTCAATTTATTTGGTCCATTCACATTTCTCAT 2820
Qy
2821 CCCTTTTCAGATAATTTAGATTTGCTTTCTTAATAAGATATTTGGAGAGCACCGTTC 2880
Db
2821 CCCTTTTCAGATAATTTAGATTTGCTTTCTTAATAAGATATTTGGAGAGCACCGTTC 2880
Qy
2881 TTATTTCAGTATTAATACTCGTCTTCTTAAGCATCCCTCAATCTTTTAATAACAATTA 2940
Db
2881 TTATTTCAGTATTAATACTCGTCTTCTTAAGCATCCCTCAATCTTTTAATAACAATTA 2940
Qy
2941 TAGCATCTAATCTTCAACAACTGGCCGTTTGTGTGAACACTCTCTTTAATAAATAATTT 3000
Db
2941 TAGCATCTAATCTTCAACAACTGGCCGTTTGTGTGAACACTCTCTTTAATAAATAATTT 3000
Qy
3001 TTCCTTTCCCAATTTCCATATAGAAATCCATCTTTCATCGCTTTTTCGTC 3060
Db
3001 TTCCTTTCCCAATTTCCATATAGAAATCCATCTTTCATCGCTTTTTCGTC 3060
Qy
3061 TCATCTGATGAATCAAAATCGCTTCTTCTGTGTCATCAAGGTTTAATTTTTTATGTA 3120
Db
3061 TCATCTGATGAATCAAAATCGCTTCTTCTGTGTCATCAAGGTTTAATTTTTTATGTA 3120

Qy 3121 TCTTTTAAACCAACCATAGAGATTAAACCTTTTACGGTGAACCTTCTCTCCAAATCA 3180
Db
3121 TCTTTTAAACCAACCATAGAGATTAAACCTTTTACGGTGAACCTTCTCTCCAAATCA 3180
Qy
3181 GACAAACGTTTCAAAATCTTTTCTTCATCATCGTCAATAAAATCCGTATCTTTACAGGA 3240
Db
3181 GACAAACGTTTCAAAATCTTTTCTTCATCATCGTCAATAAAATCCGTATCTTTACAGGA 3240
Qy
3241 TATTTTGAGTTTCTGTCGAATTCGCGATTTGATATCCGATTTATATTTATTTTTCGGTATT 3300
Db
3241 TATTTTGAGTTTCTGTCGAATTTGCGATTTGATATCCGATTTATATTTATTTTTCGGTATT 3300
Qy
3301 TTTTATTTAAACGCTCAAAATCGTTTCTGGGACGTTTTCAGGTTTATTTTCGTTTAGTTA 3360
Db
3301 TTTTATTTAAACGCTCAAAATCGTTTCTGGGACGTTTTCAGGTTTATTTTCGTTTAGTTA 3360
Qy
3361 TCGGCATTAATCGTTTAAACAGCGTTATCGTAGCTTAAAGCCCTTGAGCGTAGCGTCT 3420
Db
3361 TCGGCATTAATCGTTTAAACAGCGTTATCGTAGCTTAAAGCCCTTGAGCGTAGCGTCT 3420
Qy
3421 TTGACGCGAAGATGTTGCTGTAGATTATGAAGCCGATGACTGAATGAATAATAAAGC 3480
Db
3421 TTGACGCGAAGATGTTGCTGTAGATTATGAAGCCGATGACTGAATGAATAATAAAGC 3480
Qy
3481 GCAGCGTCTCTTCTATTTTCGTTTGGAGAGGCTCAAGCGAGTTTGAGGAAATGAATTC 3540
Db
3481 GCAGCGTCTCTTCTATTTTCGTTTGGAGAGGCTCAAGCGAGTTTGAGGAAATGAATTC 3540
Qy
3541 TCATCGGTTTGAATTTTAAATTTGCTTGCATTTTTCGCGAGCGTAGCGCTGGAATTT 3600
Db
3541 TCATCGGTTTGAATTTTAAATTTGCTTGCATTTTTCGCGAGCGTAGCGCTGGAATTT 3600
Qy
3601 TTGAAACAAATTTTGAATTTTGAACAAATTTGGAACAAATTTGGAACAAATTTTTCCT 3660
Db
3601 TTGAAACAAATTTTGAATTTTGAACAAATTTGGAACAAATTTTTCCT 3660
Qy
3661 ACTACGACCCCATTAAGTCGCGATGTCCTTGTGTCGCAAAACGCTCTATCCCA 3720
Db
3661 ACTACGACCCCATTAAGTCGCGATGTCCTTGTGTCGCAAAACGCTCTATCCCA 3720
Qy
3721 CTGGCTCAAGGTTTGAAGGTTTTCATCCCAACGATCCCAACGTTTTCGCAAC 3780
Db
3721 CTGGCTCAAGGTTTGAAGGTTTTCATCCCAACGATCCCAACGTTTTCGCAAC 3780
Qy
3781 GTTTTTTATAAATCTATATTTAAGTAGCTTTATTTGTTTATTTTATGATTACAAAGTGATA 3840
Db
3781 GTTTTTTATAAATCTATATTTAAGTAGCTTTATTTGTTTATTTTATGATTACAAAGTGATA 3840
Qy
3841 CACTAATTTTATAAATTTTGAATTTGATTTGATTTTAAATGCTGATTTTACAGATCGAAA 3900
Db
3841 CACTAATTTTATAAATTTTGAATTTGATTTTAAATGCTGATTTTACAGATCGAAA 3900
Qy
3901 AAAGAGTTATGATTTCTCTGACAAAGAGCAAGATAAATAAATAAAGATATGGCAAC 3960
Db
3901 AAAGAGTTATGATTTCTCTGACAAAGAGCAAGATAAATAAATAAAGATATGGCAAC 3960
Qy
3961 AAAAGGTTTTCACAAATCTGCGGTTGCGGTTAGCTATAGAAGATATGCAAGAAAG 4020
Db
3961 AAAAGGTTTTCACAAATCTGCGGTTGCGGTTAGCTATAGAAGATATGCAAGAAAG 4020
Qy
4021 AATCAGAACAAACAAATTAAGCGAAGCTCGGTTTAAAGAGGATACGAGTTTTCGCTA 4080
Db
4021 AATCAGAACAAACAAATTAAGCGAAGCTCGGTTTAAAGAGGATACGAGTTTTCGCTA 4080
Qy
4081 CTTGTTTTTGAAGGATAATATCATGCTATTAATAAATACTAAGCTAGAAAATTTGG 4140
Db
4081 CTTGTTTTTGAAGGATAATATCATGCTATTAATAAATACTAAGCTAGAAAATTTGG 4140
Qy
4141 ATTTTATTTATATCTCTGACTCAATTTCTTAATCGAAGAAATTAAGAGTTTGG 4200
Db
4141 ATTTTATTTATATCTCTGACTCAATTTCTTAATCGAAGAAATTAAGAGTTTGG 4200


```
QY 1671 GTTTGTTGAACTAATGGGTGCTTTAGTGAAGATAAAGACCAATTAATAAATATGGTTC 1730
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 5868
QY 7007 GTTTGTTGAACTAATGGGTGCTTTAGTGAAGATAAAGACCAATTAATAAATATGGTTC 6948
QY 1731 TTTTGTGTTTTTTTAAAGATTTAGCGTAGCGAATAATCCTTTCTTCTTCTTCTTCTGAT 1790
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 5808
QY 6947 TTTTGTGTTTTTTTAAAGATTTAGCGTAGCGAATAATCCTTTCTTCTTCTTCTTCTGAT 6888
QY 1791 AATAAGGGTAACATTTCCGGGATAGACTGTAAATCTCAGCATATAAATCCCTTTTCA 1850
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 5748
QY 6887 AATAAGGGTAACATTTCCGGGATAGACTGTAAATCTCAGCATATAAATCCCTTTTCA 6828
QY 1851 TTTTCTAATGTAATCTATTACCTTATTAATTAATCAATTCGCTCAATAATTAATCCTTTT 1910
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 5688
QY 6827 TTTTCTAATGTAATCTATTACCTTATTAATTAATCAATTCGCTCAATAATTAATCCTTTT 6768
QY 1911 TCTTATTACGCAAAATGCGCGATTTAAGCACACCCCTTTATTCGGTTAATGCGCCATGAC 1970
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 3050
QY 6767 TCTTATTACGCAAAATGCGCGATTTAAGCACACCCCTTTATTCGGTTAATGCGCCATGAC 6708
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 5628
QY 1971 AGCCATGATAATTAATACTAGGAGAAGTTAATAAATACGTAACCAACATGATTAACA 2030
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 3170
QY 6707 AGCCATGATAATTAATACTAGGAGAAGTTAATAAATACGTAACCAACATGATTAACA 6648
QY 2031 ATTATTAGAGTCAATCGTCAAAATGATGATGCGTTTTTGACACATCCCACTATATCGTG 2090
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 5508
QY 6647 ATTATTAGAGTCAATCGTCAAAATGATGATGCGTTTTTGACACATCCCACTATATCGTG 6588
QY 2091 TCGTTCTGTCACCTCCGATCCCAATCCAGAAATCTCTAGGGAATCCAGAAATTTCTC 2150
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 3230
QY 6587 TCGTTCTGTCACCTCCGATCCCAATCCAGAAATCTCTAGGGAATCCAGAAATTTCTC 6528
QY 2151 AGAGTCGGAAGTTGACACAGACATTAAGAACTGGCAGAGATGTCATAAAGCAAGGAAG 2210
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 5448
QY 6527 AGAGTCGGAAGTTGACACAGACATTAAGAACTGGCAGAGATGTCATAAAGCAAGGAAG 6468
QY 2211 ATCTGATTGCTTAATCTGCTTACAGTTAAGACCGAAGCGCTCGTGTATAACAGATGCGATG 2270
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 3290
QY 6467 ATCTGATTGCTTAATCTGCTTACAGTTAAGACCGAAGCGCTCGTGTATAACAGATGCGATG 6408
QY 2271 ATGAGACCAATCAACATGCGCACTGCAATGCTGACAGTACAGTCAAGGATGGTAGAAA 2330
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 3290
QY 6407 ATGAGACCAATCAACATGCGCACTGCAATGCTGACAGTACAGTCAAGGATGGTAGAAA 6348
QY 2331 TGTGTCGTCCTGTCACACCAATATTACGCCATTTACGCCATTTGCGTCAATTAACAGCTCTTC 2390
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 5448
QY 6347 TGTGTCGTCCTGTCACACCAATATTACGCCATTTGCGTCAATTAACAGCTCTTC 6288
QY 2391 TACGATAAGGCGCACAAATCGCATCGTGGAAAGCTTTGGGCTTTCTACCGATTAGCAGTTTG 2450
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 3290
QY 6287 TACGATAAGGCGCACAAATCGCATCGTGGAAAGCTTTGGGCTTTCTACCGATTAGCAGTTTG 6228
QY 2451 ATACACTTTCTTAAGTATCCACTGAATCATTAATCGGCAATAGAGAAATTCAGCC 2510
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 5448
QY 6227 ATACACTTTCTTAAGTATCCACTGAATCATTAATCGGCAATAGAGAAATTCAGCC 6168
QY 2511 ATGTGTAAGCGGCAATCTGATTCACCTGAGATGCATTAATCTAGTAGAATCTCTTCGCT 2570
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 5448
QY 6167 ATGTGTAAGCGGCAATCTGATTCACCTGAGATGCATTAATCTAGTAGAATCTCTTCGCT 6108
QY 2571 ATCAAAATTCATCTCCACTTCCACTCACCGGTTGTCATTCATGGCTGAATCTCTGCTTC 2630
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 5448
QY 6107 ATCAAAATTCATCTCCACTTCCACTCACCGGTTGTCATTCATGGCTGAATCTCTGCTTC 6048
QY 2631 CTCTGTTGACATGACACACATCATCTCAATATCCGAATAGGCGCCATCAGTCTCAGCACC 2690
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 5448
QY 6047 CTCTGTTGACATGACACACATCATCTCAATATCCGAATAGGCGCCATCAGTCTCAGCACC 5988
QY 2691 AAGAGAGCCATAAACCACCAATAGCGTTAAACATATCCCAATATTATCCAAATATTCGTTTC 2750
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 5448
QY 5987 AAGAGAGCCATAAACCACCAATAGCGTTAAACATATCCCAATATTATCCAAATATTCGTTTC 5928
QY 2751 CTTAAATTCATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 2810
```

```
DB 5927 CTTAAATTCATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 5868
QY 2811 TATTTCTCATTCCTCTTTCAGATAATTTTAGATTTTCTTCTTCTTCTTCTTCTTCTTCTGAG 2870
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 5808
QY 5867 TATTTCTCATTCCTCTTTCAGATAATTTTAGATTTTCTTCTTCTTCTTCTTCTTCTTCTGAG 5808
QY 2871 AGCACCGTTCTTATTCAGCTATTAATAACTCGTCTTCTTAAGCATCTCTTCAATCCTTTTA 2930
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 5748
QY 5807 AGCACCGTTCTTATTCAGCTATTAATAACTCGTCTTCTTAAGCATCTCTTCAATCCTTTTA 5748
QY 2931 ATAAACAATATAGCATCTAAATCTTCAACAACATGCGCCGTTTGTGTAACACTACTCTTTAAT 2990
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 5688
QY 5747 ATAAACAATATAGCATCTAAATCTTCAACAACATGCGCCGTTTGTGTAACACTACTCTTTAAT 5688
QY 2991 AAAATAAATTTTTCCGTTCCCAATTCACATTCGCAATTAAGAAAATCATCTTCAATCGGC 3050
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 5628
QY 5687 AAAATAAATTTTTCCGTTCCCAATTCGCAATTCGCAATTAAGAAAATCATCTTCAATCGGC 5628
QY 3051 TTTTTCGTCATCATCTGTATGATCAAAATCGCTTCTTCTGTGTGCATCAAGGTTTAAATTT 3110
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 5568
QY 5627 TTTTTCGTCATCATCTGTATGATCAAAATCGCTTCTTCTGTGTGCATCAAGGTTTAAATTT 5568
QY 3111 TTTATGTATTTCTTTTAAACAAACCAATAGGAGATTAACCTTTTACGGTGTAAACCTTTC 3170
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 5508
QY 5567 TTTATGTATTTCTTTTAAACAAACCAATAGGAGATTAACCTTTTACGGTGTAAACCTTTC 5508
QY 3171 CTCCAAATCAGACAAACGTTTCAAAATCTTCTTCTCATCATCGGTCAATAAATCCGATC 3230
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 5448
QY 5507 CTCCAAATCAGACAAACGTTTCAAAATCTTCTTCTCATCATCGGTCAATAAATCCGATC 5448
QY 3231 CTTTACAGATATTTTGCAGATTTCTCAATTCGCGATTGTATATCCGATTATATTATT 3290
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 5388
QY 5447 CTTTACAGATATTTTGCAGATTTCTCAATTCGCGATTGTATATCCGATTATATTATT 5388
QY 3291 TTTTCGTTATTTTATTATAA 3310
DB |||||||CATGAACAAATCTTCTTCTCTAGTCATTAATATTGGTCCATTCCAC 5387
QY 5387 TTTTCGTTATTTTATTATAA 5368
```

RESULT 3
AR018032/c

```
LOCUS 8119 bp DNA linear PAT 05-DEC-1998
DEFINITION Sequence 1 from patent US 5780285.
ACCESSION AR018032
VERSION AR018032.1 GI:3973635
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 8119)
AUTHORS Ballinger,M.D. and Wells,J.A.
TITLE Subtilisin variants capable of cleaving substrates containing
dibasic residues
JOURNAL Patent: US 5780285-A 1 14-JUL-1998;
FEATURES
source Location/Qualifiers
1..8119
/mol_type="unassigned DNA"
```

ORIGIN

```
Query Match 48.0%; Score 2467.2; DB 6; Length 8119;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2472; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 831 TTAGAAAACCGACTGTAAAAGTCAGTCGGCATTTATCTCATATTATAAAGCCAGTCAT 890
DB 7847 TTAGAAAACCGACTGTAAAAGTCAGTCGGCATTTATCTCATATTATAAAGCCAGTCAT 7788
QY 891 TAGGCGTATCTGCAATTCCTGAATAGAGTTTCATAAACAATCCTGTCATGATCAATCAC 950
DB 7787 TAGGCGTATCTGCAATTCCTGAATAGAGTTTCATAAACAATCCTGTCATGATCAATCAC 7728
```


CDS	complement (5365..6063) /gene="repA" /codon_start=1 /transl_table=11 /product="replication protein RepA" /protein_id="AAD46053.1" /db_xref="GI:5688892" /translations="MAIKNTKARNFGLFLYPSIDPNDWKEKLESIGVNAVSPHLHMD EKDKDTWNSDVIRNGHYKPHYHVIYIARNPVIESVRNKIRKLGNSVAHVBI LDYIKGSYELTHESDAJAKNKHIDYKDIININDFDIDRYITLDESQRKLKULL DIVDDYLVNTKOLMAFIRLGAEFGLINTNDYKIDIVSTNSSAFRLWPEGNYCCGYRA SYAKVLDAETGEIK" 7156..7215 /codon_start=1 /transl_table=11 /product="emrC leader peptide" /protein_id="AAD46054.1" /db_xref="GI:5668893" /translation="MGIFSFIVISTVHYQPNKK" 7275..8009 /gene="emrC" 7275..8009 /codon_start=1 /transl_table=11 /product="adenine methylase" /protein_id="AAD46055.1" /db_xref="GI:5668894" /translations="MNEKNIKHSQNFITSKHNDIKIMWIRLNEHNDIPEISGKGHP TLBLVKRCNFVTAIEIDHKLCKTENKLVNDHDFQVINKDIDWQFKPKNOSYKIGNI PYNISDIDIRKIVPDSIANBIYIIVRYGFAKRLNLTKRSLALLMAEVDISLSMVP EYHPKPKVNSLIRLRSKKSRIKSHKQKQKYNFVNMKNVKEYKFIITKNQFNLSLKH AGIDLLNLSFEQFLSIFNSYKLFNK"	
gene		
CDS		
Query Match	31.9%; Score 1641.6; DB 12; Length 8082;	
Best Local Similarity	98.4%; Pred. No. 1e-263;	
Matches 1689; Conservative	0; Mismatches 24; Indels 3; Gaps 3;	
QY	3279	TTTATATTTATTTTCGGTATTTTATTAAGCGTCTCAAAATCGTTTCTGGGACGTTT 3338
DB	6892	TATATATTTATTTATTCGGATTTTTTATTAAGCGTCTCAAAATCGTTTCTGGGACGTTT 6833
QY	3339	TAGCGTTTATTTTCGTTAGTTATCGGCATATCGTTAAAAACAGCGCTATCGTAGCGTAA 3398
DB	6832	TAGCGTTTATTTTCGTTAGTTATCGGCATATCGTTAAAAACAGCGCTATCGTAGCGTAA 6773
QY	3399	AAGCCCTTGACGCTAGCGT-GCTTTGACGGAAGATGTTGCTGTAGATTATGAAGCC 3457
DB	6772	AAGCCCTTGACGCTAGCGTGGCTTTGACGGAAGATGTTGCTGTAGATTATGAAGCC 6713
QY	3458	GATGACTGAATGAATAATAGCGCAGCGTCTTCTATTTTCGGTTGGAGGAGGCTCAAGG 3517
DB	6712	GATGACTGAATGAATAATAGCGCAGCGCTTCTATTTTCGGTTGGAGGAGGCTCAAGG 6653
QY	3518	GAGTTTGAGGAATGAATAATCCCTCATCGGTTTGATTTAAAAATTCGTTGCAATTTTCG 3577
DB	6652	GAGTTTGAGGAATGAATAATCCCTCATCGGTTTGATTTAAAAATTCGTTGCAATTTTCG 6593
QY	3578	CGAGCGGTAGCGCTGGAATAATTTTGAATAATTTTGAATAATTTTGAATAATTTTGAATA 3637
DB	6592	CGAGCGGTAGCGCTGGAATAATTTTGAATAATTTTGAATAATTTTGAATAATTTTGAATA 6533
QY	3638	AAGGAAGCGAATTTTTCGTTCCGTACTACGACCCCCCAATTAAGTCCCGAGTGCCAAATTTT 3697
DB	6532	AAGGAAGCGAATTTTTCGTTCCGTACTACGACCCCCCAATTAAGTCCCGAGTGCCAAATTTT 6473
QY	3698	GTGCCAAAACGCTCTATCCCACTGCGCTCAAGGGTTGAGGGTTTTCGATCCGAGTCCCAAC 3757
DB	6472	GTGCCAAAACGCTCTATCCCACTGCGCTCAAGGGTTTTCGATCCGAGTCCCAAC 6413
QY	3758	GAATGCCCAACGTTTTCGCCCAACGTTTTCGATCCCACTGCGCTCAAGGGTTTTCGATCCG 3817

ORIGIN

DB	6412	GAATCGCCCAAGTTTTTCGCCCAACGTTTTTATATAAATCTATATTTAAGTAGCTTTATTGTT 6353
QY	3818	GTTTTTATGATTACAAAGTGATACACTAAATTTTATATAAATTTATTGATTGAGTTTATTA 3877
DB	6352	GTTTTTATGATTACAAAGTGATACACTAACTTTTATAAATTTATTGATTGAGTTTATTA 6293
QY	3878	AATGTTGATTTTCAGAAATCGAAAAAAGAGTTATGATTTCTCTGACAAAAGAGCAAGATAA 3937
DB	6292	AATGTTGATTTTCAGAAATCGAAAAAAGAGTTATGATTTCTCTGACAAAAGAGCAAGATAA 6233
QY	3938	AAAAATTAACAGATATGCGGAAACAAAAGGTTTTTCAAAATCTCGGTTGCGGCTTAGC 3997
DB	6232	AAAAATTAACAGATATGCGGAAACAAAAGGTTTTTCAAAATCTCGGTTGCGGCTTAGC 6173
QY	3998	TATAGAACAATATGCAAGAAAGGAATCAGAAACAAAATAAAGCGAAAGCTCGCGTTTT 4057
DB	6172	TATAGAACAATATGCAAGAAAGGAATCAGAAACAAAATAAAGCGAAAGCTCGCGTTTT 6113
QY	4058	TAGAAGGATACGAGTTTTTCGTTACTGTTTGTGTTTGAATAAGGTAA-TATATCATGGCTATTAA 4116
DB	6112	TAGAAGGATACGAGTTTTTCGTTACTGTTTGTGTTTGAATAAGGTAA-TATATCATGGCTATTAA 6053
QY	4117	AAATACATAAGCTAGAAATTTTGGATTTTATTTATATCTCTGACTCAATTCCTATGATTG 4176
DB	6052	AAATACATAAGCTAGAAATTTTGGATTTTATTTATATCTCTGACTCAATTCCTATGATTG 5993
QY	4177	GAAGAAAAATTTAGAGAGTTTTGGGCGTATCTATGGCTGTCTAGTCTCTTTACACGATATGA 4236
DB	5992	GAAGAAAAATTTAGAGAGTTTTGGGCGTATCTATGGCTGTCTAGTCTCTTTACACGATATGA 5933
QY	4237	CGAAAAAAGATAAAGATACATGGAATAGTAGTGTATTTACGAAATGGAAGACACTA 4296
DB	5932	CGAAAAAAGATAAAGATACATGGAATAGTAGTGTATTTACGAAATGGAAGACACTA 5873
QY	4297	TAAAAAACACACTATACGTTTATATATTTGACGAAATCTGTAAACAATAGAAAGCGT 4356
DB	5872	TAAAAAACACACTATACGTTTATATATTTGACGAAATCTGTAAACAATAGAAAGCGT 5813
QY	4357	TAGGAACAAGATTAAGCGAAAAATTTGGGAAATAGTTTCACTCATGATTTTGAAGATCTTGA 4416
DB	5812	TAGGAACAAGATTAAGCGAAAAATTTGGGAAATAGTTTCACTCATGATTTTGAAGATCTTGA 5753
QY	4417	TTATATCAAGGTTTATATGATTTTTCGCTCATGATTTTGAAGATCTTGAAGATCTTGA 4476
DB	5752	TTATATCAAGGTTTATATGATTTTTCGCTCATGATTTTGAAGATCTTGAAGATCTTGA 5693
QY	4477	TAAACATATATACGCAAAAAAGATTTTGAACATTAATGATTTTGAATTTGATTTGACCGCTA 4536
DB	5692	TAAACATATATACGCAAAAAAGATTTTGAACATTAATGATTTTGAATTTGATTTGACCGCTA 5633
QY	4537	TATAACACTTTGATGAAGCAAAAAAGAGATTTGAAGATTTTGAAGATTTTGAAGATTTGAT 4596
DB	5632	TATAACACTTTGATGAAGCAAAAAAGAGATTTGAAGATTTTGAAGATTTTGAAGATTTGAT 5573
QY	4597	TGACTATATTTTGGTAAATACAAAAAGATTTTAAATGCGCTTTTATTCGCCCTTAGGGGACGGA 4656
DB	5572	TGACTATATTTTGGTAAATACAAAAAGATTTTAAATGCGCTTTTATTCGCCCTTAGGGGACGGA 5513
QY	4657	GTTTGGAAATTTTAAATACGAAATGATGTAAGAGATTTGTTTCAACAAACTCTAGCGCCTT 4716
DB	5512	GTTTGGAAATTTTAAATACGAAATGATGTAAGAGATTTGTTTCAACAAACTCTAGCGCCTT 5453
QY	4717	TAGATTTGCTTTGAGGCAATTTATCAGTGTGATATAGAGCAAGTTATGCAAAAGGTTCT 4776
DB	5452	TAGATTTGCTTTGAGGCAATTTATCAGTGTGATATAGAGCAAGTTATGCAAAAGGTTCT 5393
QY	4777	TGATGCTGAAACGGGGGAAATAAAAATGACAAAACAAAGAAAAAGAGTTATTTGCTGAAAT 4836
DB	5392	TGATGCTGAAACGGGGGAAATAAAAATGACAAAACAAAGAAAAAGAGTTATTTGCTGAAAT 5333
QY	4837	GAGGAATTTAAAAAGAAATTTAAGGCTTTAAAGGCGGTATTGAAAGATACAGAGAAATG 4896
DB	5332	GAGGAATTTAAAAAGAAATTTAAGGCTTTAAAGGCGGTATTGAAAGATACAGAGAAATG 5273

Qy 4897 GAAGTTGAATTAAGTACAAATAGATTATTGAGAGGAGGATTATTGAATAAATAAAA 4956
 Db 5272 GAAGTTGAATTAAGTACAAATAGATTATTGAGAGGAGGATTATTGAATAAATAAAA 5213

Qy 4957 GCCCCCTGACGAAGTCGGACTTCGTTCTTTT 4992
 Db 5212 G-CCCCCTGACGAAGTCGAAGCGGGTTCGTTCTTTT 5178

RESULT 8
 AF076213/c 8142 bp DNA linear SYN 08-FEB-2000
 LOCUS AF076213 Promoter screenings vector pM225, complete sequence.
 DEFINITION AF076213 Promoter screenings vector pM225
 ACCESSION AF076213
 VERSION AF076213.1 GI:5668895
 SOURCE Promoter screenings vector pM225
 ORGANISM Promoter screenings vector pM225
 REFERENCE 1 (bases 1 to 8142)
 AUTHORS Vriesema,A.J., Brinkman,R., Kok,J., Dankert,J. and Zaat,S.A.
 TITLE Broad-host-range shuttle vectors for screening of regulated promoter activity in viridans group streptococci: isolation of a pH-regulated promoter
 JOURNAL Appl. Environ. Microbiol. 66 (2), 535-542 (2000)
 MEDLINE 20120502
 PUBMED 10653715
 REFERENCE 2 (bases 1 to 8142)
 AUTHORS Vriesema,A.J.M., Brinkman,R., Kok,J., Dankert,J. and Zaat,S.A.J.
 TITLE Direct Submission
 JOURNAL Submitted (02-JUL-1998) Medical Microbiology, Academic Medical Center, University of Amsterdam, Meibergdreef 15, Amsterdam 1105 AZ, The Netherlands

FEATURES
 source
 1. 8142
 /organism="Promoter screenings vector pM225"
 /mol_type="genomic DNA"
 /db_xref="taxon:79020"
 26..68
 189..198
 212..1006
 /gene="aph-III"
 /note="promoterless"
 212..1006
 /gene="aph-III"
 /note="APH-III"
 /codon_start=1
 /transl_table=11
 /product="3', 5'-aminoglycoside phosphotransferase type III"
 /protein_id="AAD46056.1"
 /db_xref="GI:5668896"
 /translations="MAKMRISPELKKLIEKRCVKDTEGMSPAKVYKLVGENENLYLK MTPSRVKTVDYVEREKDMMLEGLPVKCVLHPRHDPGWSNLLMSEADGVLCEEY EDQSQKGIIEIYACIRLPHSIDSCPTNSLDRLELDYLLNLDLADVDCEWE EDTFDPKRLYDFLTKPEELVFSHGLDGSNIFVKGKVSFGIDLGRGRADKW YDIAFCVRSIREDBEQVVEFLDGLLGKPDWEKIKYIILDELF"
 1030..4520
 /note="trp-lacZ fusion"
 1254..1257
 1265..4339
 /gene="lacZ"
 /note="promoterless"
 1265..4339
 /gene="lacZ"
 /codon_start=1
 /transl_table=11
 /product="beta-d-galactosidase"
 /protein_id="AAD46057.1"
 /db_xref="GI:5668897"
 /translations="MTMIDTSLAVIQORDWENPQVQTNRLAAHPFPFAWRNSBEAR TDRPSQQLRSLNGEWRFPFAPFVPSWLECDLPEADTVVVVPSNQMHGYDAPIYT

Qy 3279 TTTATATTTATTTTCGGTATTTTATTAACGTCCTCAAAATCGTTCTCGGACGCTTT 3338
 Db 6952 TATATATTTATTTATTCGGATTTTATTAACGTCCTCAAAATCGTTCTCGGACGCTTT 6993
 Qy 3339 TAGCGTTTATTTTCGTTAGTTATCGGCATATCGTTAAAAACAGCGCTTATCGTAGCGTAA 3398
 Db 6992 TAGCGTTTATTTTCGTTAGTTATCGGCATATCGTTAAAAACAGCGCTTATCGTAGCGTAA 6933
 Qy 3399 AAGCCCTTGACGTAGCGT-GCTTTGACGGAAGATGTTCTCTGTAGATTATGAAGCC 3457
 Db 6832 AAGCCCTTGACGTAGCGTCTTTGACGGAAGATGTTCTCTGTAGATTATGAAGCC 6773
 Qy 3458 GATGACTGAATGAATAATAAGCGCAGCGCTCTCTTCTATTTTCGGTTGAGGAGGCTCAAGG 3517
 Db 6772 GATGACTGAATGAATAATAAGCGCAGCGCTCTCTTCTATTTTCGGTTGAGGAGGCTCAAGG 6713

ORIGIN
 Query Match 31.9%; Score 1641.6; DB 12; Length 8142;
 Best Local Similarity 98.4%; Pred. No. 1e-263;
 Matches 1689; Conservative 0; Mismatches 24; Indels 3; Gaps 3;

gene
 CDS
 complement (5425..6123)
 /gene="repA"
 /complement (5425..6123)
 /gene="repA"
 /codon_start=1
 /transl_table=11
 /product="replication protein RepA"
 /protein_id="AAD46058.1"
 /db_xref="GI:5668898"
 /translation="MAIKNTKARNFQFLYPDSIPNDWKEKLESGLVSNVAPSLHMD EKDKTWNSSDVIRNGKHVYKPHYHVIYIARNPVTIESVRNKIKKLGNSVLAHVEI LDYIKGYSYETHESKDALAKKHIDYKDIILINDIDRIDYITLDESQKRELKLL DLVDDNLANVTKDLMAFIRLGAEGFLINTNDVKDIVSTNSSAFRLWFEQYQCGYRA SYAKVLDATGEIK"
 7216..7275
 /codon_start=1
 /transl_table=11
 /product="emrC leader peptide"
 /protein_id="AAD46059.1"
 /db_xref="GI:5668899"
 /translation="MGIPSTFVISTVHYQPNKK"
 7335..8069
 /gene="emrC"
 7335..8069
 /gene="emrC"
 /codon_start=1
 /transl_table=11
 /product="adenine methylase"
 /protein_id="AAD46060.1"
 /db_xref="GI:5668900"
 /translation="MNEKNIKHSQNFITSKHNDIKMTNIRLNEHNIPEIGSKGHP TLELVKRCNFVTAIEDHKLCKTENKLVHDNDFOVLNKDIMOFPKQSKYKIYNI PYNISTDIIRKIVFDSIANEIVLIVEYGFARLLNTRSLALLMAEVIDISILSMVPR EYFHPKPKVNSLIRLSRKKSRISHKDKQKYNFVWKVNKEYKKIFTKQNFNSLKH AGIDDLNINISPEQFLSFPNSYKLFNK"

QY 3518 GAGTTTGAGGGATGAATTCCTCATGGGTTTGATTTTAAAAATTGCTTGCAATTTCG 3577
DB 6712 GAGTATGAGGGATGAATTCCTCATGGGTTTGATTTTAAAAATTGCTTGCAATTTCG 6653
QY 3578 CGAGCGGTAGCGCTGAAAAATTTTGAATAAATTTGGAATTTGGAATAAATTTGGGGGA 3637
DB 6652 CGAGCGGTAGCGCTGAAAAATTTTGAATAAATTTGGAATTTGGAATAAATTTGGGGGA 6593
QY 3638 AAGGAAGCAATTTTCTTCCTGCTACTACGACCCCAATTAAGTCGCGAGTGCCAAATTTT 3697
DB 6592 AAGGAAGCAATTTTCTTCCTGCTACTACGACCCCAATTAAGTCGCGAGTGCCAAATTTT 6533
QY 3698 GTCCCAAAAACGCTATCCCACTCGCTCAAGGGTTTGAAGGGTTTCAATTCGCCAAC 3757
DB 6532 GTCCCAAAAACGCTATCCCACTCGCTCAAGGGTTTGAAGGGTTTCAATTCGCCAAC 6473
QY 3758 GAATCGCCAAAGTTTTCGCCAAAGTTTATTAATCTATATTAAGTAGCTTTATGTT 3817
DB 6472 GAATCGCCAAAGTTTTCGCCAAAGTTTATTAATCTATATTAAGTAGCTTTATGTT 6413
QY 3818 GTTTTATGATTTACAAAGTACACTAAATTTTATAAATTTATTTGATTTGAGTTTATTA 3877
DB 6412 GTTTTATGATTTACAAAGTACACTAAATTTTATAAATTTATTTGATTTGAGTTTATTA 6353
QY 3878 AATGGTGATTTCAAGTTCGAAAAAAGAGTTATGATTTCTCTGACAAAAGAGCAATTA 3937
DB 6352 AATGGTGATTTCAAGTTCGAAAAAAGAGTTATGATTTCTCTGACAAAAGAGCAATTA 6293
QY 3938 AATATTAACAGATATCGCGAAACAAAAAGGTTTTCAAAATCTGCGGTTGCGCGTTAGC 3997
DB 6292 AATATTAACAGATATCGCGAAACAAAAAGGTTTTCAAAATCTGCGGTTGCGCGTTAGC 6233
QY 3998 TATAGAAGATATCGAAGAAAGGAATCAGAACAAAAAATAAAGCAAGAGCTCGCGTTT 4057
DB 6232 TATAGAAGATATCGAAGAAAGGAATCAGAACAAAAAATAAAGCAAGAGCTCGCGTTT 6173
QY 4058 TAGAAGATACGAGTTTTCCTACTCTGTTTTCATAGGTAA-TATATCATGCTATTA 4116
DB 6172 TAGAAGATACGAGTTTTCCTACTCTGTTTTCATAGGTAA-TATATCATGCTATTA 6113
QY 4117 AATATCTAAAGCTAGAAATTTTGATTTTATATATCTGACTCAATTCCTAATGATTG 4176
DB 6112 AATATCTAAAGCTAGAAATTTTGATTTTATATATCTGACTCAATTCCTAATGATTG 6053
QY 4177 GAAAGAAAAATAGAGAGTTTGGCGTACTATGCTGCTGAGTCTTACACATATGA 4236
DB 6052 GAAAGAAAAATAGAGAGTTTGGCGTACTATGCTGCTGAGTCTTACACATATGA 5993
QY 4237 CGAAAAAAGATTAAGATACATCGAATAGTAGTGTATATCGAATCGAAAGCACTA 4296
DB 5992 CGAAAAAAGATTAAGATACATCGAATAGTAGTGTATATCGAATCGAAAGCACTA 5933
QY 4297 TAAAAAACCACTATACGTTATATATATTCACGAAATCCCTGTAACAATAGAACGCT 4356
DB 5932 TAAAAAACCACTATACGTTATATATATTCACGAAATCCCTGTAACAATAGAACGCT 5873
QY 4357 TAGAAGCAAGATTAAGCAAAATTTGGGAATAGTGTCTCATCTGAGATCTTGA 4416
DB 5872 TAGAAGCAAGATTAAGCAAAATTTGGGAATAGTGTCTCATCTGAGATCTTGA 5813
QY 4417 TTATATCAAGGTTTCATATGAATTTTGACTCATGAATCAAGGAGCCTATTGTAAGAA 4476
DB 5812 TTATATCAAGGTTTCATATGAATTTTGACTCATGAATCAAGGAGCCTATTGTAAGAA 5753
QY 4477 TAAACATATACGCAAAAAAGATATTTTGAACATTAATGATTTTATATGACCGCTA 4536
DB 5752 TAAACATATACGCAAAAAAGATATTTTGAACATTAATGATTTTATATGACCGCTA 5693
QY 4537 TATACACTTGTATGAAGCCAAAAAGAGATTTGAAGAAATTTATCTTTAGATATAGTGA 4596
DB 5692 TATACACTTGTATGAAGCCAAAAAGAGATTTGAAGAAATTTATCTTTAGATATAGTGA 5633
QY 4597 TGACTATAATTTGGTAAATACAAAAGATTTAATGCTTTTATTCGCTTTAGGGAGCGGA 4656

DB 5632 TGACTATAATTTGGTAAATACAAAAGATTTAATGGCTTTTATTGCGCTTAGGGAGCGGA 5573
QY 4657 GTTTGGAAATTTAAATACGAATGATGATAAAGATATTGTTTCAACAAACTCTAGCGCTT 4716
DB 5572 GTTTGGAAATTTAAATACGAATGATGATAAAGATATTGTTTCAACAAACTCTAGCGCTT 5513
QY 4717 TAGATTATGTTTTCAGGGCAATTAATCATGTGTGGATATAGAGCAAGTTATGCAAAAGTTCT 4776
DB 5512 TAGATTATGTTTTCAGGGCAATTAATCATGTGTGGATATAGAGCAAGTTATGCAAAAGTTCT 5453
QY 4777 TGATGCTGAACCGGGGGAATAAATGACAAAACAAAGAAAAGAGTTTATTGCTGAAAAAT 4836
DB 5452 TGATGCTGAACCGGGGGAATAAATGACAAAACAAAGAAAAGAGTTTATTGCTGAAAAAT 5393
QY 4837 CAGGAATTTAAAAAAGAAATTAAGGACTTAAAGAGCGTATTGAAAGATACAGAGAAATG 4896
DB 5392 CAGGAATTTAAAAAAGAAATTAAGGACTTAAAGAGCGTATTGAAAGATACAGAGAAATG 5333
QY 4897 GAAGTTGAATTAAGTACAACTAGATTTATTGAGAGGAGGATTTTGAATAATAAATAA 4956
DB 5332 GAAGTTGAATTAAGTACAACTAGATTTATTGAGAGGAGGATTTTGAATAATAAATAA 5273
QY 4957 GCGCCCTCGACGAAAGTCGCGACTTCGTTCTTTTT 4992
DB 5272 G-CGCTGACGAAAGTCGAGGGGTTTTTATTTT 5238

RESULT 9
LLREPA L.lactis plasmid pWV01 2178 bp DNA linear BCT 18-OCT-1991
LOCUS ORF8 B, C, and D.
DEFINITION
ACCESSION X58954.1 GI:44058
VERSION repA gene; replication protein; replication protein A.
KEYWORDS Lactococcus lactis
SOURCE Lactococcus lactis
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.

REFERENCE 1
AUTHORS Leenhouts,K.J., Tolner,B., Bron,S., Kok,J., Venema,G. and Seegers,J.F.
TITLE Nucleotide sequence and characterization of the broad-host-range lactococcal plasmid pWV01
JOURNAL Plasmid 26 (1), 55-66 (1991)
MEDLINE 92052683
PUBMED 1840693
REFERENCE 2 (bases 1 to 2178)
AUTHORS Leenhouts,K.J.
TITLE Direct Submission
JOURNAL Submitted (10-DEC-1990) K.J. Leenhouts, INSTITUTE OF GENETICS, UNIVERSITY OF GRONINGEN, KERKLAAN 30, 9751 NN HAREN, THE NETHERLANDS

FEATURES
Source
1..2178
/organism="Lactococcus lactis"
/mol_type="genomic DNA"
/db_xref="taxon:1358"
/plasmid="pWV01"
355..544
/gene="ORF B"
335..544
/gene="ORF B"
/codon_start=1
/transl_table=11
/protein_id="CAA40270.1"
/db_xref="GI:44059"
/db_xref="GOA:Q48693"
/translation="MGKEANFASVLRPIKCRVFVFKTLYPNWLKLRGFSSTANE STFGPTFFINLYLSFSFVFMITK"
522..550

promoter 522..550


```
Qy 4796 ATAAATGACAAACAAGAAAAAGAGTTATTTCTGAAATAGAGGAATTAATAAAAAAGAAA 4855
Db 1503 ATAAATGACAAACAAGAAAAAGAGTTATTTCTGAAATAGAGGAATTAATAAAAAAGAAA 1562
Qy 4856 TTAAGGACTTAAAGAGCGTATTGAAGATACAGAGAAATGGAAGTTGAATTAAAGTACAA 4915
Db 1563 TTAAGGACTTAAAGAGCGTATTGAAGATACAGAGAAATGGAAGTTGAATTAAAGTACAA 1622
Qy 4916 CAATAGATTATTGAGAGGAGGATTATTGAATAAATAAAGCCCTTGACGAAAGTCG 4975
Db 1623 CAATAGATTATTGAGAGGAGGATTATTGAATAAATAAAG-CCCCTTGACGAAAGTCG 1681
Qy 4976 CGACTTCGTTCTTTT 4992
Db 1682 AAGGGGGTTTTTATTTT 1698

RESULT 10
AF372620 5126 bp DNA circular SYN 31-MAY-2002
LOCUS GlnQ allelic exchange vector pAG101, complete sequence.
ACCESSION AF372620
VERSION AF372620.1 GI:14211230
KEYWORDS
SOURCE glnQ allelic exchange vector pAG101
ORGANISM GlnQ allelic exchange vector pAG101
REFERENCE Tamura,G.S., Nittayajarn,A. and Schoentag,D.J.
AUTHORS A glutamine transport gene, glnQ, is required for fibronectin
TITLE adherence and virulence of group B streptococci
JOURNAL Infect. Immun. 70 (6), 2877-2885 (2002)
MEDLINE 22006888
PUBMED 12010975
REFERENCE Tamura,G.S., Schoentag,D.I. and Nittayajarn,A.
AUTHORS Direct Submission
TITLE Submitted (20-APR-2001) Department of Pediatrics, Division of
JOURNAL Infectious Diseases, Children's Hospital and Regional Medical
Center and the University of Washington, 4800 Sand Point Way NE,
Seattle, WA 98105-0371, USA
FEATURES
source
1. .5126
/organism="glnQ allelic exchange vector pAG101"
/mol_type="genomic DNA"
/db_xref="taxon:159742"
/lab_host="Streptococcus agalactiae"
/plasmid="pAG101"
/notes="temperature-sensitive plasmid"
order(4747..5126,859..1116)
/notes="glnQ gene interrupted by erm gene"
102..839
/gene="erm"
102..839
/gene="erm"
/notes="confers erythromycin-resistance"
/codon_start=1
/transl_table=11
/product="ribosomal RNA methylase"
/protein_id="AAK57385.1"
/db_xref="GI:14211231"
/translation="MNFKIKYSQNFLSEKVLNQI IKQNLKETDVTVEIGTKGHLT
TKLAKISQVTSIEDSHLENLSEKULKNTVLIHQDILQFQFNKQRIKVGSIIP
YHLSTQIKKVVFEASHADIIYIEEGFYKRLDIDHRLGLLHTQVSIQOOLKJFAB
CFHPKPKVNSYLIIKLRHTDIDVPDKYWKLYTFYFVKWNVREYQLFTKQFHQAMKHA
KVNLSVTYVQVLSIFNSYLLFNKRK"
3845..4495
/gene="cat"
3845..4495
/gene="cat"
/notes="confers chloramphenicol resistance"
/codon_start=1
```

```
/transl_table=11
/product="chloramphenicol acetyl transferase"
/protein_id="AAK57386.1"
/db_xref="GI:14211232"
/translation="MNFKIKIDLDNWKREIPIHNYLNOOTPTSITEIDISVLYNNIKQ
EGYFYFAFIVLRVNTSNTATGYNSDELGYWDKLEPLYTFPGVSTFGSIWT
PVKNDFKFVLDLYLSDVEKNGSKLPKTPIPENAFSLSIIPWTSFTGNLNNNS
NYLLPIITAGNFINKNSIYIPLSLQVHHSVCDGYHAGLFMNSIQELSDRDNWLL"

ORIGIN
Query Match 31.8%; Score 1635.2; DB 12; Length 5126;
Best Local Similarity 98.2%; Pred. No. 1.3e-262;
Matches 1685; Conservative 0; Mismatches 28; Indels 3; Gaps 3;

Qy 3279 TTTATATTTATTTTCGGTATTTTATTAAACGCTCTCAAAATCGTTCTCTGGACGTTT 3338
Db 1395 TATATATATTTATTTATCCGATTTTATTATAAAGCTCTCAAAATCGTTCTGAGACGTTT 1454
Qy 3339 TAGCGTTTATTTTCGTTTATTCGGCATATCGTATAAAGAGCGGTTATCGTAGCGTAA 3398
Db 1455 TAGCGTTTATTTTCGTTTATTCGGCATATCGTATAAAGAGCGGTTATCGTAGCGTAA 1514
Qy 3399 AAGCCCTTGAGCGTAGCGT-GCTTTGCGAGGAGATGTTGCTGTAGATTATGAAGCC 3457
Db 1515 AAGCCCTTGAGCGTAGCGTGTTCAGCGAGAGATGTTGCTGTAGATTATGAAGCC 1574
Qy 3458 GATGACTGAATGAAATAATAAGCGCAGCGCTCTCTTATTTCGGTTGGAGGCTCAAGG 3517
Db 1575 GATGACTGAATGAAATAATAAGCGCAGCGCTCTTATTTCGGTTGGAGGCTCAAGG 1634
Qy 3518 GAGTTTGAGGGAATGAAATTCCTCATCGGTTTGAATTTTAAATTTGCTTGCATTTTGC 3577
Db 1635 GAGTATGAGGGAATGAAATTCCTCATCGGTTTGAATTTTAAATTTGCTTGCATTTTGC 1694
Qy 3578 CGAGCGGTAGCGCTCGAAATTTTGAATAAATTTGGAATTTGAAAAAATTCGGGGGA 3637
Db 1695 CGAGCGGTAGCGCTCGAAATTTTGAATAAATTTGGAATTTGAAAAAATTCGGGGGA 1754
Qy 3638 AAGGAAGCGAATTTTTCGCTTCCGTAACAGACCCCATTAAGTCCGAGTGCCATTTT 3697
Db 1755 AAGGAAGCGAATTTTTCGCTTCCGTAACAGACCCCATTAAGTCCGAGTGCCATTTT 1814
Qy 3698 GTGCCAAAACCGCTCTATCCCACTGGCTCAAGGGTTTGAAGGGTTTCAATCCCAAC 3757
Db 1815 GTGCCAAAACCGCTCTATCCCACTGGCTCAAGGGTTTGAAGGGTTTCAATCCCAAC 1874
Qy 3758 GAATCGCCAAACGTTTTCGCCAACGTTTTTATATACTATATTTAAGTAGCTTTATGTT 3817
Db 1875 GAATCGCCAAACGTTTTCGCCAACGTTTTTATATACTATATTTAAGTAGCTTTATGTT 1934
Qy 3818 GTTTTATGATTACAAAGTGATACACTAATTTTATAAATTTATTTGATGGATTTTATA 3877
Db 1935 GTTTTATGATTACAAAGTGATACACTAATTTTATAAATTTATTTGATGGATTTTATA 1994
Qy 3878 AATGGTGATTTTCAGAAATCGAAAAAGAGTTATGATTTCTCTGACAAAAGAGCAAGATAA 3937
Db 1995 AATGGTGATTTTCAGAAATCGAAAAAGAGTTATGATTTCTCTGACAAAAGAGCAAGATAA 2054
Qy 3938 AAAATTAACAGATATGGCGAAACAAAAAGTTTTCAAAAATCTCGGTTGCGCGCTTAGC 3997
Db 2055 AAAATTAACAGATATGGCGAAACAAAAAGTTTTCAAAAATCTCGGTTGCGCGCTTAGC 2114
Qy 3998 TATAGAAGATATGCAAGAAAGGAATCAGAACAAAAAATAAGCGAAAGCTCGCGTTT 4057
Db 2115 TATAGAAGATATGCAAGAAAGGAATCAGAACAAAAAATAAGCGAAAGCTCGCGTTT 2174
Qy 4058 TAGAAGGATACGAGTTTTCGCTACTTGTGTTTGAAGAAGTAA-TATATCATCGGTATTAA 4116
Db 2175 TAGAAGGATACGAGTTTTCGCTACTTGTGTTTGAAGAAGTAA-TATATCATCGGTATTAA 2234
Qy 4117 AAATACTAAGCTAGAAATTTTGGATTTTATATATCTGCTGACTCAATTCCTAATGATTG 4176
Db 2235 AAATACTAAGCTAGAAATTTTGGATTTTATATATCTGCTGACTCAATTCCTAATGATTG 2294
```


/db_xref="GI:3127120"
/translation="MWISSEKRVMLSLTKEQKXKLTMDAKOKGFSKAVPALAIBEY
ARKSEQKX"

ORIGIN

Query Match	31.8%;	Score	1635.2;	DB 12;	Length	9021;	
Best Local Similarity	98.2%;	Pred.	No. 1.2e-262;				
Mismatches	0;	Conservative	28;	Indels	3;	Gaps	3;
Matches	1685;						
QY	3279	TTTATATTTATTTTTCGGTATTTTATTAATAAACGTCCTCAAAATCGTTTCTGGGAGCTTT	3338				
DB	8804	TATATATTTATTTATTCGGATTTTATTAATAAACGTCCTCAAAATCGTTTCTGAGACGTTT	8745				
QY	3339	TAGCGTTTATTTTCGTTAGTTATCGGCATATCGTTAAACAGCGCTATCGTAGCGTAA	3398				
DB	8744	TAGCGTTTATTTTCGTTAGTTATCGGCATATCGTTAAACAGCGCTATCGTAGCGTAA	8685				
QY	3399	AAGCCCTTGAGCGTAGCGT-GCTTTGCGAGCAAGATGTTGTCTGTGTAGATTATGAAGCC	3457				
DB	8684	AAGCCCTTGAGCGTAGCGTGGCTTTGCGAGCAAGATGTTGTCTGTGTAGATTATGAAGCC	8625				
QY	3458	GATGACTGAATGAATAATAAGCGCAGCGCTCTCTATTTTCGTTGCGAGGCGCTCAAGG	3517				
DB	8624	GATGACTGAATGAATAATAAGCGCAGCGCTCTCTATTTTCGTTGCGAGGCGCTCAAGG	8565				
QY	3518	GAGTTTCAGGGAATGAATTCCTCATGCGTTTGATTTTAAAAATTCCTTGCAATTTTCG	3577				
DB	8564	GAGTTTCAGGGAATGAATTCCTCATGCGTTTGATTTTAAAAATTCCTTGCAATTTTCG	8505				
QY	3578	CGAGCGTAGCGCTGGAATAATTTTGAATAAATTTTGGAAATTTTGGAAAAAATTTGGGGGA	3637				
DB	8504	CGAGCGTAGCGCTGGAATAATTTTGAATAAATTTTGGAAATTTTGGAAATTTTGGAAAAAATTTGGGGGA	8445				
QY	3638	AAGGAACGCAATTTTGTCTCCGTACTACGACCCCGCTTAAAGTCCGAGTGCCTCAATTTT	3697				
DB	8444	AAGGAACGCAATTTTGTCTCCGTACTACGACCCCGCTTAAAGTCCGAGTGCCTCAATTTT	8385				
QY	3698	GTGCCAAAAACGCTCTATCCCACTGGCTCAAGGGTTTGAGGGGTTTTCATTCGCCAAC	3757				
DB	8384	GTGCCAAAAACGCTCTATCCCACTGGCTCAAGGGTTTGAGGGGTTTTCATTCGCCAAC	8325				
QY	3758	GAATCGCGCAACGTTTTCGCCAACGTTTTTATAATCTATATTAAAGTACGCTTATGTT	3817				
DB	8324	GAATCGCGCAACGTTTTCGCCAACGTTTTTATAATCTATATTAAAGTACGCTTATGTT	8265				
QY	3818	GTTTTTATGATTTACAAAGTGATACACTAAATTTTATAAAATTTTGAATTCGAGTTT	3877				
DB	8264	GTTTTTATGATTTACAAAGTGATACACTAAATTTTATAAAATTTTGAATTCGAGTTT	8205				
QY	3878	AATGGTGATTTTCAAGATCGAAAAAGATGATGATTTCTGCAAAAAGAGCAAGATAA	3937				
DB	8204	AATGGTGATTTTCAAGATCGAAAAAGATGATGATTTCTGCAAAAAGAGCAAGATAA	8145				
QY	3938	AAATTAACAGATATGCGCAACAAAGGTTTTTCAAAATCTCGGTTGCGGCTTAGC	3997				
DB	8144	AAATTAACAGATATGCGCAACAAAGGTTTTTCAAAATCTCGGTTGCGGCTTAGC	8085				
QY	3998	TATAGAAGATATCAAGAAAGGATCAGAACAAAAAATAAGCGAAAGCTCCGCTTT	4057				
DB	8084	TATAGAAGATATCAAGAAAGGATCAGAACAAAAAATAAGCGAAAGCTCCGCTTT	8025				
QY	4058	TAGAAGGATACAGTTTTTCGCTACTCTGTTTTGTATAGGTAA-TATATCATGGCTATTAA	4116				
DB	8024	TAGAAGGATACAGTTTTTCGCTACTCTGTTTTGTATAGGTAA-TATATCATGGCTATTAA	7965				
QY	4117	AAATACCTAAAGCTAGAAATTTTGGATTTTATATATCTCTGACTCAATTCCTAATGATG	4176				
DB	7964	AAATACCTAAAGCTAGAAATTTTGGATTTTATATATCTCTGACTCAATTCCTAATGATG	7905				
QY	4177	GAAGAAAAATPAGAGGTTTGGCGGTATCTATGCGCTGATGCTTTTACAGTATGGA	4236				
DB	7904	GAAGAAAAATPAGAGGTTTGGCGGTATCTATGCGCTGATGCTTTTACAGTATGGA	7845				

QY	4237	CGAAAAAAGATAAAGATACATGGAATAGTAGTGATGTTATACGAAATCGAAAGCACTA	4296				
DB	7844	CGAAAAAAGATAAAGATACATGGAATAGTAGTGATGTTATACGAAATCGAAAGCACTA	7785				
QY	4297	TAAAAAACCCACATATACAGTTATATATATGACGAAATCCTTACAAATAGAAAGCGT	4356				
DB	7784	TAAAAAACCCACATATACAGTTATATATATGACGAAATCCTTACAAATAGAAAGCGT	7725				
QY	4357	TAGGAACAAGATTAAGCGAAATTTGGGGAATAGTTCAAGTTGCTCATGTTGAGATCTTGA	4416				
DB	7724	TAGGAACAAGATTAAGCGAAATTTGGGGAATAGTTCAAGTTGCTCATGTTGAGATCTTGA	7665				
QY	4417	TTATATCAAGGTTTCATATGAATATTTGACCTCATGAATCAAAGGACGCTATTTCTAGAA	4476				
DB	7664	TTATATCAAGGTTTCATATGAATATTTGACCTCATGAATCAAAGGACGCTATTTCTAGAA	7605				
QY	4477	TAAACATATATACGACAAAGAAATTTTGAACATTAATGATTTGATTTGACCCGCTA	4536				
DB	7604	TAAACATATATACGACAAAGAAATTTTGAACATTAATGATTTGATTTGACCCGCTA	7545				
QY	4537	TATACACTTGTATGAAGCCAAAAAGAGAAATTTCAAGAAATTTACTTTTATAGATATAGTGA	4596				
DB	7544	TATACACTTGTATGAAGCCAAAAAGAGAAATTTCAAGAAATTTACTTTTATAGATATAGTGA	7485				
QY	4597	TGACTATATTTTGGTAAATACAAAGATTTAATGCTTTTATTCGCTTAGGGAGCGGA	4656				
DB	7484	TGACTATATTTTGGTAAATACAAAGATTTAATGCTTTTATTCGCTTAGGGAGCGGA	7425				
QY	4657	GTTCGGAATTTTAAATACGAATGATTAAGAGATTTGTTTCAACAAACTCTAGCGCTT	4716				
DB	7424	GTTCGGAATTTTAAATACGAATGATTAAGAGATTTGTTTCAACAAACTCTAGCGCTT	7365				
QY	4717	TAGATTTAGTTTGGTGGGCAATATATCAGTTGGATATAGCAAGTATGCAAGTTCT	4776				
DB	7364	TAGATTTAGTTTGGTGGGCAATATATCAGTTGGATATAGCAAGTATGCAAGTTCT	7305				
QY	4777	TCATGCTGAACCGGGGAAATTAATAATGACAAAGAAAGAGTATTTTCTGCAAAAT	4836				
DB	7304	TCATGCTGAACCGGGGAAATTAATAATGACAAAGAAAGAGTATTTTCTGCAAAAT	7245				
QY	4837	GAGGAATTTAAAAAAGAAATTAAGGACTTAAAGAGCGTATTGAAAGATACAGAGAAATG	4896				
DB	7244	GAGGAATTTAAAAAAGAAATTAAGGACTTAAAGAGCGTATTGAAAGATACAGAGAAATG	7185				
QY	4897	GAGGTTGAATTAAGTACACATAGATTTTATGAGGAGGAGGATTTTGAATAATAAAA	4956				
DB	7184	GAGGTTGAATTAAGTACACATAGATTTTATGAGGAGGAGGATTTTGAATAATAAAA	7125				
QY	4957	GCCCCCTCGCAAAAGTCGCGACTTCGTTCTTTTTT 4992					
DB	7124	G-CCCCCTCGCAAAAGTCGAGGGGTTTTTTTTT 7090					

RESULT 12
AY028776/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

AY028776 13561 bp DNA circular SYN 05-APR-2001
TnphoZ mutagenesis vector pMHL120, complete plasmid sequence.
AY028776
AY028776.1 GI:13549398
TnphoZ mutagenesis vector pMHL120
TnphoZ mutagenesis vector pMHL120
artificial sequences; vectors.
1 (bases 1 to 13561)
Lee M.H., Nittayajarn A., Ross R.P., Rothschild C.B., Parsons D.,
Claborn A., and Rubens C.E.
Characterization of Enterococcus faecalis alkaline phosphatase and
use in identifying Streptococcus agalactiae secreted proteins
J. Bacteriol. 181 (18), 5790-5799 (1999)
99412280
10482522
2 (bases 1 to 13561)
Lee M.H., Clancy A., Jones A.L., Nittayajarn A. and Rubens C.E.

QY	3938	AAAAATTAACAGATATGCGGAAACAAAAAGTTTTC	3997	RESULT 13	LLREPAB
Db	12684	AAAAATTAACAGATATGCGGAAACAAAAAGTTTTC	12625	LOCUS	Lactococcus lactis plasmid replicon copx and repX genes.
QY	3998	TATAGAAGATATGCAAGAAAGAAATCAGAACAAAAA	4057	DEFINITION	2505 bp DNA linear BCT 20-MAY-1998
Db	12624	TATAGAAGATATGCAAGAAAGAAATCAGAACAAAAA	12565	ACCESSION	X54310.1 GI:3150197
QY	4058	TAGAAGATATGCAAGTTTTCGCTACTTGTGTTTTC	4116	VERSION	copx gene; replication protein; repX gene.
Db	12564	TAGAAGATATGCAAGTTTTCGCTACTTGTGTTTTC	12505	KEYWORDS	Lactococcus lactis
QY	4117	AAATTAACAGATATGCGGAAACAAAAAGTTTTC	4176	SOURCE	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
Db	12504	AAATTAACAGATATGCGGAAACAAAAAGTTTTC	12445	REFERENCE	1
QY	4177	GAAGAAATATGAGAGTTTTCGCGTATCTATGCGCT	4236	AUTHORS	Xu, F.F., Pearce, L.E. and Yu, P.L.
Db	12444	GAAGAAATATGAGAGTTTTCGCGTATCTATGCGCT	12385	TITLE	Genetic analysis of a lactococcal plasmid replicon
QY	4237	CGAAAAAAGATATAAGATACATGGAATAGTATGAT	4296	JOURNAL	Mol. Gen. Genet. 227 (1), 33-39 (1991)
Db	12384	CGAAAAAAGATATAAGATACATGGAATAGTATGAT	12325	MEDLINE	91260676
QY	4297	TAAAAACCCACATATCAGTTATATATGTCAGCGAAT	4356	PUBMED	1904536
Db	12324	TAAAAACCCACATATCAGTTATATATGTCAGCGAAT	12265	REFERENCE	2
QY	4357	TAGAACACAGATTAAGCGAAATTTGGGGAATAGTT	4416	AUTHORS	Grohmann, E., Moscoso, M., Zechner, E.L., Espinosa, M. and del Solar, G.
Db	12264	TAGAACACAGATTAAGCGAAATTTGGGGAATAGTT	12205	TITLE	In vivo definition of the functional origin of leading strand
QY	4417	TTATATCAAAAGTTTCAATGAAATTTTGAATCTAT	4476	JOURNAL	Unpublished
Db	12204	TTATATCAAAAGTTTCAATGAAATTTTGAATCTAT	12145	REFERENCE	3
QY	4477	TAAACATATATACGACAAAAAGATATTTGACATTA	4536	AUTHORS	Yu, P.L.
Db	12144	TAAACATATATACGACAAAAAGATATTTGACATTA	12085	TITLE	Direct Submission
QY	4537	TATAACACTTGTATGAAAGCCAAAAAGAGAAATTT	4596	JOURNAL	Submitted (18-MAY-1998) Espinosa M., Centro de Investigaciones
Db	12084	TATAACACTTGTATGAAAGCCAAAAAGAGAAATTT	12025	COMMENT	Biologicas, CSIC, Velazquez, 144. E-28006, Madrid, SPAIN
QY	4597	TGACTATAATTTGTTAAATCAAAAGATTTTAAATG	11965	FEATURES	On May 22, 1998 this sequence version replaced gi:44063.
Db	12024	TGACTATAATTTGTTAAATCAAAAGATTTTAAATG	11905	repeat_region	Location/Qualifiers
QY	4657	GTTTGGAAATTTTAAATACGAATGATGTAAGATAT	11964	rep origin	1. .2505
Db	11964	GTTTGGAAATTTTAAATACGAATGATGTAAGATAT	4716	repeat_region	/organism="Lactococcus lactis"
QY	4717	TAGATTATGTTTTCAGGCAATTTATCAGTGTGGAT	4776	-35_signal	/mol_type="genomic DNA"
Db	11904	TAGATTATGTTTTCAGGCAATTTATCAGTGTGGAT	11845	-10_signal	/strain="ssp. lactis 5136"
QY	4777	TGATGCTCAACCGGGGAATTAATATGCAACAAAG	4836	gene	/db_xref="taxon:1358"
Db	11844	TGATGCTCAACCGGGGAATTAATATGCAACAAAG	11785	CDS	/plasmid="pFX2"
QY	4837	GAGGAATTTAAAAAGAAATTAAGSACTTAAAGAG	4896	attenuator	1054. .1093
Db	11784	GAGGAATTTAAAAAGAAATTAAGSACTTAAAGAG	11725	terminator	1087. .1155
QY	4897	GAAATTGAAATTAAGTACAAATAGATTATTGAG	4956	RBS	1211. .1248
Db	11724	GAAATTGAAATTAAGTACAAATAGATTATTGAG	11665	gene	1289. .1994
QY	4957	GCCCCCTTGACGAAAGTCGCGACTTCGTTCTTTT	4992	CDS	1312. .1317
Db	11664	GCCCCCTTGACGAAAGTCGCGACTTCGTTCTTTT	11630		1329. .1332

```

/protein_id="CAA38208.1"
/db_xref="GI:3150198"
/db_xref="GOA:Q04138"
/db_xref="SPRMBL:Q04138"
/translation="MAIKNTKARNFGFLLYPPDISPNDWKEKLESIGVSNVAPSLHMD
EKDKDWTMSDYIRNGKHYPKHVHYIARNPVYIESVRNKRKIKLGNSSVAHVEI
LDYIKGSYEYLDKAKNKHIDILNINDPDIIDRYITLDESOKRELKILL
DIVDDYLVNTKDMAFILRLGAEPGILNTNDVKDIVTNSAFRLWPEGNQCYRA
SYAKVLDAETGEIK"

```

ORIGIN

```

Query Match      31.7%; Score 1629.6; DB 1; Length 2505;
Best Local Similarity 98.4%; Pred. No. 1.2e-261;
Matches 1688; Conservative 0; Mismatches 24; Indels 4; Gaps 4;

3279 TTTATATTTTTCGGTATTTTATTAATAAAGCTCTCAAAATCGTTCTGGAGCGTTT 3338
Dbb|||||
741 TATATATTTTATTCGCGATTTTATTAATAAAGCTCTCAAAATCGTTCTGGAGCGTTT 800
Dbb|||||
3339 TAGCGTTTATTCGTTAGTTATCGGCATAATCGTTAAACAGCGGTTATPCSPAGCGTAA 3398
Qyb|||||
801 TAGCGTTTATTCGTTAGTTATCGGCATAATCGTTAAACAGCGGTTATPCSPAGCGTAA 860
Dbb|||||
3399 AAGCCCTTGAGCGTAGCGT-GCTTTCGACGAGATGTTCTGTTAGATTTATGAAGCC 3457
Qyb|||||
861 AAGCCCTTGAGCGTAGCGTGGCTTTTCGACGAGATGTTCTGTTAGATTTATGAAGCC 920
Dbb|||||
3458 GATGACTGAATGAATAATAAGCGCAGCGTCTTCTATTTTCGTTGGAGGAGGCTCAAGG 3517
Qyb|||||
921 GATGACTGAATGAATAATAAGCGCAGCGTCTTCTATTTTCGTTGGAGGAGGCTCAAGG 980
Dbb|||||
3518 GAGTTGAGGAGATGAATTCCTCATCGGTTTGATTTTAAATAATTCCTTGCATTTTGC 3577
Qyb|||||
981 GAGTATGAGGAGATGAATTCCTCATCGGTTTGATTTTAAATAATTCCTTGCATTTTGC 1040
Dbb|||||
3578 CGACGCTAGCGCTGGAAAAATTTTGAATAAATTTTGGAAATTTGAAAAAATTTGGGGGA 3637
Dbb|||||
1041 CGACGCTAGCGCTGGAAAAATTTTGAATAAATTTTGGAAATTTGAAAAAATTTGGGGGA 1100
Dbb|||||
3638 AAGGAACGAAATTTGCTTCGTAACGACCCCGCATTAAGTTCGAGTGCACATTTTTC 3697
Qyb|||||
1101 AAGGAACGAAATTTGCTTCGTAACGACCCCGCATTAAGTTCGAGTGCACATTTTTC 1160
Dbb|||||
3698 GTGCCAAAACGCTCTATCCCACTCGCTCAAGGTTTCGAGGTTTTCATTCGCGCAAC 3757
Qyb|||||
1161 GTGCCAAAACGCTCTATCCCACTCGCTCAAGGTTTTCGAGGTTTTCATTCGCGCAAC 1220
Dbb|||||
3758 GAATCGCCCAACGTTTTCGCAACGTTTTCATTAATCTATTAATTAAGTACGTTTATGTT 3817
Dbb|||||
1221 GAATCGCCCAACGTTTTCGCAACG-TTTTATTAATCTATTAATTAAGTACGTTTATGTT 1279
Dbb|||||
3818 GTTTTATGATTAACAAGTACATCACTAAATTTTATTAATAATTTTGAATTTGGAGTTT 3877
Qyb|||||
1280 GTTTTATGATTAACAAGTACATCACTAAATTTTATTAATAATTTTGAATTTGGAGTTT 1339
Dbb|||||
3878 AATGGTGATTTCAAGATTCGAAAAAAGAGTTATGATTTCTCTGACAAAAGAGCAAGATAA 3937
Qyb|||||
1340 AATGGTGATTTCAAGATTCGAAAAAAGAGTTATGATTTCTCTGACAAAAGAGCAAGATAA 1399
Dbb|||||
3938 AAAATTAACAGATATGCGGAAACAAAAGGTTTTCAAAATCTGCGGTTTCGCGGTTAGC 3997
Qyb|||||
1400 AAAATTAACAGATATGCGGAAACAAAAGGTTTTCAAAATCTGCGGTTTCGCGGTTAGC 1459
Dbb|||||
3998 TATAGAGAATATGCAAGAAGGAAATCAGAACAAAAAATAAGCGAAAGCTCGCGTTT 4057
Qyb|||||
1460 TATAGAGAATATGCAAGAAGGAAATCAGAACAAAAAATAAGCGAAAGCTCGCGTTT 1519
Dbb|||||
4058 TAGAAGGATACGAGTTTTCGCTACTGTTTTCGATTAAGGTAA-TATATCATGCTATTAA 4116
Qyb|||||
1520 TAGAAGGATACGAGTTTTCGCTACTGTTTTCGATTAAGGTAAATATCATGCTATTAA 1579
Dbb|||||
4117 AAATACTAAAGCTAGAAATTTTCGATTTTATTAATATCTGACTCAATTCCTAATGATG 4176
Qyb|||||

```

```

Db 1580 AATACTAAAGCTAGAAATTTTGATTTTATATATCTCTGACTCAATTCCTAATGATTG 1639
Qyb|||||
4177 GAAAGAAAAATTTAGAGAGTTTGGCGGTATCTATGGCTGTGAGTCTCTTTACACGATATGA 4236
Dbb|||||
1640 GAAAGAAAAATTTAGAGAGTTTGGCGGTATCTATGGCTGTGAGTCTCTTTACACGATATGA 1699
Qyb|||||
4237 CGAAAAAAGATAAAGATACATGGAATAGTAGTGTATATAGAAATCGAAAGCACTA 4296
Dbb|||||
1700 CGAAAAAAGATAAAGATACATGGAATAGTAGTGTATATAGAAATCGAAAGCACTA 1759
Qyb|||||
4297 TAAARACCACTATCACGTTATATATATTGACGAAATCCCTGTACATACATGAAGCGT 4356
Dbb|||||
1760 TAAARACCACTATCACGTTATATATATTGACGAAATCCCTGTACATACATGAAGCGT 1819
Qyb|||||
4357 TAGGAACAAGATTAAAGCGAAATTTGGGGAATAGTTCAGTTGCTCATCTTGGAGATCTTGA 4416
Dbb|||||
1820 TAGGAACAAGATTAAAGCGAAATTTGGGGAATAGTTCAGTTGCTCATCTTGGAGATCTTGA 1879
Qyb|||||
4417 TTATATCAAGGTTTCATATGAATATTTGACTCATGAATCAAGGACGCTATTGCTAAGAA 4476
Dbb|||||
1880 TTATATCAAGGTTTCATATGAATATTTGACTCATGAATCAAGGACGCTATTGCTAAGAA 1939
Qyb|||||
4477 TAAACATATATACGACAAAAAAGATATTTGAACATTAATGAATTTTGATATTCACCGCTA 4536
Dbb|||||
1940 TAAACATATATACGACAAAAAAGATATTTGAACATTAATGAATTTTGATATTCACCGCTA 1999
Qyb|||||
4537 TATAACACTTGAATGAAGCCAAAAAGAGAAATTTGAAGAAATTTACTTTTAGATATAGTGA 4596
Dbb|||||
2000 TATAACACTTGAATGAAGCCAAAAAGAGAAATTTGAAGAAATTTACTTTTAGATATAGTGA 2059
Qyb|||||
4597 TGACTATAATTTTGTAAATACAAAGATTTTAATGGCTTTTATTCGCTTAGCGGAGCGGA 4656
Dbb|||||
2060 TGACTATAATTTTGTAAATACAAAGATTTTAATGGCTTTTATTCGCTTAGCGGAGCGGA 2119
Qyb|||||
4657 GTTTGGAAATTTTAAATACGAATGATTAAGAGATTTTGTTCACAAATCTTAGCGGCTT 4716
Dbb|||||
2120 GTTTGGAAATTTTAAATACGAATGATTAAGAGATTTTGTTCACAAATCTTAGCGGCTT 2179
Qyb|||||
4717 TAGATTTAGTTTTCAGCGCAATTTATCAGTGTGATATAGAGCAAGTTATTCGAAAGTTCT 4776
Dbb|||||
2180 TAGATTTAGTTTTCAGCGCAATTTATCAGTGTGATATAGAGCAAGTTATTCGAAAGTTCT 2239
Qyb|||||
4777 TGATGCTCAAAACGGGGGAAATTAATAACAAAACAAAAGAGTTTATTTGCTGAAAAAT 4836
Dbb|||||
2240 TGATGCTCAAAACGGGGGAAATTAATAACAAAACAAAAGAGTTTATTTGCTGAAAAAT 2299
Qyb|||||
4837 GAGGAATTAATAAAGAAATTTAGGACTTTAAAGAGCGTATTAAGAGATACAGAGAAATG 4896
Dbb|||||
2300 GAGGAATTAATAAAGAAATTTAGGACTTTAAAGAGCGTATTAAGAGATACAGAGAAATG 2359
Qyb|||||
4897 GAAGTTGAATTAAGTACAAATAGATTTTATGAGAGGAGGATTTATGAAATTAATAA 4956
Dbb|||||
2360 GAAGTTGAATTAAGTACAAATAGATTTTATGAGAGGAGGATTTATGAAATTAATAA 2419
Qyb|||||
4957 GCCCCCTGAGAAAGTCGCGACTTCCTCTCTTTT 4992
Dbb|||||
2420 G-CGCCCTGACAAAGTCGAAGGGGGTTTATTTT 2454

```

RESULT 14

```

AB055651
LOCUS      AB055651      4438 bp      DNA      circular SYN 21-MAY-2003
DEFINITION Thermosensitive suicide vector pSET5s DNA, complete sequence.
ACCESSION  AB055651
VERSION    AB055651.1  GI:15553035
KEYWORDS   Thermosensitive suicide vector pSET5s
SOURCE      Thermosensitive suicide vector pSET5s
ORGANISM   Thermosensitive suicide vector pSET5s
REFERENCE   1
AUTHORS     Takamatsu, D., Osaki, M. and Sekizaki, T.
TITLE       Thermosensitive suicide vectors for gene replacement in
            Streptococcus suis

```

JOURNAL Plasmid 46 (2), 140-148 (2001)
MEDLINE 21476064
PUBMED 11591139
REFERENCE 2 (bases 1 to 4438)
AUTHORS Takamatsu, D., Osaki, M. and Sekizaki, T.
TITLE Direct Submission
JOURNAL Submitted (08-FEB-2001) Daisuke Takamatsu, National Institute of
Animal Health, Laboratory of Molecular Bacteriology, Kannondai
3-1-1, Tsukuba, Ibaraki 305-0856, Japan
(E-mail: p1013d@nih.affrc.go.jp, Tel: 81-298-38-7743,
Fax: 81-298-38-7743)
FEATURES
Location/Qualifiers
1..4438
/organism="Thermosensitive suicide vector pSET5s"
/mol_type="genomic DNA"
/db_xref="taxon:152115"
528..689
/gene="orfC"
528..689
/gene="orfC"
/codon_start=1
/transl_table=11
/product="OrfC"
/protein_id="BAB64884.1"
/db_xref="GI:15553036"
/translation="MVISESKRWISLTKEQDKLTDMAKQKGFSAVALAIEY
ARXSEOKK"
756..1454
/gene="repAts"
756..1454
/gene="repAts"
/codon_start=1
/transl_table=11
/product="RepAts"
/protein_id="BAB64895.1"
/db_xref="GI:15553037"
/translation="MAIKNTKARNFGLLYPDSIPDNWKEKLSLGVSMVSPHMD
EKKDDTWNNISITQNGKHYKPHYIYIARNPVITSEVRNKKIKLGNSSVAHVEI
LDYKGSVEYLTHESDAIAKNKHIDYKDLINIDFDIDRYITLDESQKRELKLL
DIVDYNLVNTKQMAFIRLGRAEFGILNTDVKDIVSTNSSAFRLWFEGNYCGYRA
SYAKVDAETGEIK"
complement(1803..2126)
/gene="lacZ"
complement(1803..2126)
/gene="lacZ"
/codon_start=1
/transl_table=11
/product="LacZ"
/protein_id="BAB64886.1"
/db_xref="GI:15553038"
/translation="MTWITPSLHACRSTLSDPRVPSSNSLAVLQRRDWNENPGVTQLN
RLAAHPFPASWRNSEEARTRPSQQLSLNGEWMRYFLTLHLCGISHRIWCTLSTI
CSDAA"
3593..4243
/gene="cat"
3593..4243
/gene="cat"
/codon_start=1
/transl_table=11
/product="chloramphenicol acetyltransferase"
/protein_id="BAB64887.1"
/db_xref="GI:15553039"
/translation="MNFNKIDLDWKKIEFNHYLNQQTTFSTITEIDISVLNFKQ
EGYKFPAPFIFLTVRVINSNTAFRTGNSDGLGWDKLFYITFDGVSRTSGIWT
PVKDFEFDLYLSDVEKYNGSKLPKTPENAFSLSIPTWSTFGNLSINNNNS
NYLLPIITAGKFKNGSIYLPFLSLQVHVSVDGYHAGLFMNSIQELSDRPNWLL"
ORIGIN
Query Match 30.6%; Score 1575.2; DB 12; Length 4438;
Best Local Similarity 98.7%; Pred. No. 1.3e-252;
Matches 1619; Conservative 0; Mismatches 18; Indels 3; Gaps 3;
QY 3355 TAGTTATCGCATATCGTTAAACACAGCGGTATCGTAGCGTAAAGCCCTTGACGGTAG 3414

3 TAGTTATCGCATATCGTTAAACACAGCGGTATCGTAGCGTAAAGCCCTTGACGGTAG 62
3415 CQT-GCTTTGACGAGAGATGTTCTGTAGATTATGAAGCCGATGACTGAATGAAT 3473
63 CQTGGCTTTGCAGCAGAGATGTTCTGTAGATTATGAAGCCGATGACTGAATGAAT 122
3474 AATAAGCGAGCGTCTCTTCTATTTCCGTTGGAGGAGGTCAAGGAGTTTGAGGGAATGA 3533
123 AATAAGCGAGCGTCTCTTCTATTTCCGTTGGAGGAGGTCAAGGAGTATGAGGGAATGA 182
3534 AATTCCTCATCGGTTTGATTTTAAAAATGCTTCAATTTTGGCCAGCGGTAGCGCTGG 3593
183 AATTCCTCATCGGTTTGATTTTAAAAATGCTTCAATTTTGGCCAGCGGTAGCGCTGG 242
3594 AAAAAATTTGAAAAAATTTGGAATTTGAAAAAATGCGGGAAGGAAGCGAATTTTG 3653
243 AAAAAATTTGAAAAAATTTGGAATTTGAAAAAATGCGGGAAGGAAGCGAATTTTG 302
3654 CTTCCGTACTACGACCCGCCCATTAAGTGCAGAGTCCCAATTTTGTGCCAAAAACGCTCT 3713
303 CTTCCGTACTACGACCCGCCCATTAAGTGCAGAGTCCCAATTTTGTGCCAAAAACGCTCT 362
3714 ATCCCAACTGGCTCAAGGTTTGAGGGGTTTTCATGCGCAACGAATCGCCAAAGTTT 3773
363 ATCCCAACTGGCTCAAGGTTTAAAGGGTTTTCATGCGCAACGAATCGCCAAAGTTT 422
3774 CGCCAAAGTTTTCATTAATCTATATTAAGTAGTCTTATTTGTTGTTTATGATTACAA 3833
423 CGCCAAAGTTTTCATTAATCTATATTAAGTAGTCTTATTTGTTGTTTATGATTACAA 482
3834 AGTGATACACTAATTTTATAAATTTATTTGATGAGTCTTATTAATGTTGATTTCAGAA 3893
483 AGTGATACACTAATTTTATAAATTTATTTGATGAGTCTTATTAATGTTGATTTCAGAA 542
3894 TCGAAAAAAGAGTATGATTTCTCTGACAAAAGAGCAGATAAAAAATTAACAGATATG 3953
543 TCGAAAAAAGAGTATGATTTCTCTGACAAAAGAGCAGATAAAAAATTAACAGATATG 602
3954 GCGAAAAAAGAGTCTTTCAAAAATCTCGGTTGCGGTTAGCTATAGAAAGATATGCA 4013
603 GCGAAAAAAGAGTCTTTCAAAAATCTCGGTTGCGGTTAGCTATAGAAAGATATGCA 662
4014 AGAAGGATCAGAACAAAAAATAAGCGAAGTCTCGTTTTCAGAGGATAGAGTT 4073
663 AGAAGGATCAGAACAAAAAATAAGCGAAGTCTCGTTTTCAGAGGATAGAGTT 722
4074 TTGCTACTTGTCTTTGATAAGTAA-TATATCATCGTCTTATTAATACTAAAGCTAGA 4132
723 TTGCTACTTGTCTTTGATAAGTAA-TATATCATCGTCTTATTAATACTAAAGCTAGA 782
4133 AATTTGGATTTTATATATCTGACTCAATCTCTAATGATGGAAGAAAAATTAGAG 4192
783 AATTTGGATTTTATATATCTGACTCAATCTCTAATGATGGAAGAAAAATTAGAG 842
4193 AGTTTGGGCGTATCTATGCTGTCAGTCTTTTACAGATATGAGCAAAAAAAGATAAA 4252
843 AGTTTGGGCGTATCTATGCTGTCAGTCTTTTACAGATATGAGCAAAAAAAGATAAA 902
4253 GATACATGGAATAGTAGTATGATATACGAATGGAAGCACTATAAAAAACACACTAT 4312
903 GATACATGGAATAGTAGTATATATATATACGAATGGAAGCACTATAAAAAACACACTAT 962
4313 CAGCTTATATATTCGACGAATCTCTGTAACATAGAAAGCGTTAGGAACAAGATTAAG 4372
963 CAGCTTATATATTCGACGAATCTCTGTAACATAGAAAGCGTTAGGAACAAGATTAAG 1022
4373 CGAAAAATGGGGAATAGTTTCAAGTTGCTCATGTTGAGATACTTGAATATCAAAAGTTCA 4432
1023 CGAAAAATGGGGAATAGTTTCAAGTTGCTCATGTTGAGATACTTGAATATCAAAAGTTCA 1082
4433 TATGAATATTTGACTCATGAATCAAAAGCGTATTTGCTAAGNATAAACATATACGAC 4492

Db 303 CTTCGGTACTACGACCCCCCAATTAAAGTCCGAGTGCCCAATTTTGTGCGCAAAACGCTCT 362
QY 3714 ATCCCAACTGCTCAAGGGTTGAGGGTTTTCAATCGCAACGAATGCCCAACGTTTT 3773
Db 363 ATCCCAACTGCTCAAGGGTTTAAAGGGTTTTTCAATCGCCCAACGAATGCCCAACGTTTT 422
QY 3774 CGCCAAAGCTTTTATATAAATCTATATTAAAGTAGCTTTATTTGTTGTTTTTATGATTACAA 3833
Db 423 CGCCAAAGCTTTTATATAAATCTATATTAAAGTAGCTTTATTTGTTGTTTTTATGATTACAA 482
QY 3834 AGTCATACACTPAATTTTATAAATTTTGAATGAGTTTTTAAATGGTGATTTCAGAA 3893
Db 483 AGTGATACACTAATTTTATAAATTTTGAATGAGTTTTTAAATGGTGATTTCAGAA 542
QY 3894 TCGAAAAAAGAGTATGATTTCTCTGACAAAAGAGCAAGTAAAAAATTAACAGATATG 3953
Db 543 TCGAAAAAAGAGTATGATTTCTCTGACAAAAGAGCAAGTAAAAAATTAACAGATATG 602
QY 3954 GCGAAACAAAAGTTTTTCAAAATCTGCGGTTGCGGCTTAGCTATAGAAAGTAATGCA 4013
Db 603 GCGAAACAAAAGTTTTTCAAAATCTGCGGTTGCGGCTTAGCTATAGAAAGTAATGCA 662
QY 4014 AGAAAGGAATCAGAACAAAATAAAGCGAAGCTCGGTTTTTAGAAGGATACGAGTT 4073
Db 663 AGAAAGGAATCAGAACAAAATAAAGCGAAGCTCGGTTTTTAGAAGGATACGAGTT 722
QY 4074 TTGCTACTTGTGTTTGAATAGGTAA-TATATCATGCTATTAAAAATCTAAAGCTAGA 4132
Db 723 TTGCTACTTGTGTTTGAATAGGTAA-TATCATGCTATTAAAAATCTAAAGCTAGA 782
QY 4133 AATTTTGGATTTTATATATCTCGACTCAATTCCTAATGATTGGAAGAAAAAATTAGAG 4192
Db 783 AATTTTGGATTTTATATATCTCGACTCAATTCCTAATGATTGGAAGAAAAAATTAGAG 842
QY 4193 AGTTTGGCGGTATCTATGGCTGTGAGTCCCTTACACGATATGGACGAAAAAAGATAAA 4252
Db 843 AGTTTGGCGGTATCTATGGCTGTGAGTCCCTTACACGATATGGACGAAAAAAGATAAA 902
QY 4253 GATACATGGAATAGTAGTGTGTTATACGAAATGGAAAGCACTATATAAAACCCACACTAT 4312
Db 903 GATACATGGAATATAGTGTGTTATATATACAAATGGAAAGCACTATATAAAACCCACACTAT 962
QY 4313 CACGTTATATATGACGAAATCCCTGTAAATAGAAAGCTTAGGAAACAAGTTAAG 4372
Db 963 CACGTTATATATGACGAAATCCCTGTAAATAGAAAGCTTAGGAAACAAGTTAAG 1022
QY 4373 CGAAAATGGGCAATAGTTCAGTTGCTCATGTTGAGATCTTGATTATATCAAAAGTTCA 4432
Db 1023 CGAAAATGGGCAATAGTTCAGTTGCTCATGTTGAGATCTTGATTATATCAAAAGTTCA 1082
QY 4433 TATGAATATTTGACTCATGAATCAAAAGACGCTATTGCTAAGAAATAAACATATATACGAC 4492
Db 1083 TATGAATATTTGACTCATGAATCAAAAGACGCTATTGCTAAGAAATAAACATATATACGAC 1142
QY 4493 AAAAAAGATATTTTGAACATTAAATGATTGATATGACCGCTATATAACACTTTGATGAA 4552
Db 1143 AAAAAAGATATTTTGAACATTAAATGATTGATATGACCGCTATATAACACTTTGATGAA 1202
QY 4553 AGCCAAAAAGAGAAATGAAGAATTTACTTTTAGATATAGTGGATGACTATAATTTGGTA 4612
Db 1203 AGCCAAAAAGAGAAATGAAGAATTTACTTTTAGATATAGTGGATGACTATAATTTGGTA 1262
QY 4613 AATACAAAAGATTTAATGGCTTTTATTCGCCCTTAGGGAGCGGAGTTTGGAAATTTAAAT 4672
Db 1263 AATACAAAAGATTTAATGGCTTTTATTCGCCCTTAGGGAGCGGAGTTTGGAAATTTAAAT 1322
QY 4673 ACGAATGATATAAGATATTTGTTTCAACAACTCTAGCGCTTTAGATTATGTTTGGAG 4732
Db 1323 ACGAATGATATAAGATATTTGTTTCAACAACTCTAGCGCTTTAGATTATGTTTGGAG 1382
QY 4733 GCGAATTAATCATGTGGATATAGAGCAAGTTATGCAAGGTTCTTGATGCTGAAACCGGG 4792

Db 1383 GCGAATTTATCAGTGTGGATATAGACCAAGTTATGCAAAAGTTCTTGATGCTGAAACGGGG 1442
QY 4793 GGAATATAAATGACAAACAAAGAAAGAGTTATTTGCTGAAATGAGGAATTAABAAAAG 4852
Db 1443 GGAATATAAATGACAAACAAAGAAAGAGTTATTTGCTGAAATGAGGAATTAABAAAAG 1502
QY 4853 AATTTAAAGGACTTAAAGAGCGCTATTGAAAGATACAGAGAAATGGAAGTTGAAATTAAGTA 4912
Db 1503 AATTTAAAGGACTTAAAGAGCGCTATTGAAAGATACAGAGAAATGGAAGTTGAAATTAAGTA 1562
QY 4913 CAACAATAGATTATTGAGAGGAGGATTTGAATATAAAGCCGCCCTGACGAAAG 4972
Db 1563 CAACAATAGATTATTGAGAGGAGGATTTGAATATAAAG-CCCCCTGACGAAAG 1621
QY 4973 TCGCGACTTCGTTCTTTTTT 4992
Db 1622 TCGAAGGGGGTTTTTTATTTT 1641

Search completed: September 24, 2004, 04:19:41
Job time : 12732.2 secs

This Page Blank (uspio)